


```

RN NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Jorde M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Letrelle P., Layman D., Ozersky P., Rohlfing T.,
RA Schert P., Walker K., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Marais E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanaki J.P.,
RA Peyrard M., Kedra D., Serousi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22."
RL Nature 402:489-495(1999).
[7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 191-360 (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirogawa M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N.,
RA Ohara O.;
RT "Characterization of cDNA clones selected by the GeneMark analysis
RT from size-fractionated cDNA libraries from human brain."
RL DNA Res. 6:329-336(1999).
CC -!- FUNCTION: Carrier protein. Binds to some hydrophobic molecules and
CC promotes their transfer between the different cellular sites.
CC Binds with high affinity to alpha-tocopherol. Also binds with a
CC weaker affinity to other tocopherols and to tocotrienols. May have
CC a transcriptional activator activity via its association with
CC alpha-tocopherol. Probably recognizes and binds some squalene
CC structure, suggesting that it may regulate cholesterol
CC biosynthesis by increasing the transfer of squalene to a metabolic
CC active pool in the cell.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic in absence of alpha-tocopherol,
CC and nuclear in presence of alpha-tocopherol.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1;
CC IsoId=O76054-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O76054-2; Sequence=VSP_006031;
CC
CC Name=3;
CC IsoId=O76054-3; Sequence=Not described;
CC -!- TISSUE SPECIFICITY: Widely expressed. Strong expression in liver,
CC brain and prostate.
CC -!- DEVELOPMENTAL STAGE: Low expression in fetal tissues.
CC -!- SIMILARITY: Contains 1 CRAL-TRIO domain.
CC -!- SIMILARITY: Contains 1 GOLD domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AL096881; CAB51405.1; -; mRNA.
CC EMBL; CR456571; CAG30457.1; -; mRNA.
CC EMBL; AC004832; AAF19256.1; -; Genomic_DNA.
CC EMBL; AB033012; BAA86500.1; -; mRNA.
CC PIR; JC7708; JC7708
CC
CC PDB; 1O6U; X-ray; A/C/B=1-403.
CC PDB; 1OLM; X-ray; A/C/B=1-403.
CC Ensembl; ENSG00000100003; Homo sapiens.
CC HGNC; HGNC:10699; SEC14L2.
CC MIM; 607558; -;
CC GO; GO:0005737; Cytoplasm; NAS.
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0005386; F:carrier activity; NAS.
CC GO; GO:0005543; F:phospholipid binding; NAS.
CC GO; GO:0016563; F:transcriptional activator activity; NAS.
CC GO; GO:0008431; F:vitamin E binding; NAS.
CC GO; GO:0045893; P:positive regulation of transcription, DNA-d. .; NAS.
CC GO; GO:0045544; P:regulation of cholesterol biosynthesis; NAS.
CC InterPro; IPR001071; CRAL_bd_toc_tran.
CC InterPro; IPR001251; CRAL_bd_TRIO_C.
CC InterPro; IPR008273; CRAL_bd_TRIO_N.
CC InterPro; IPR000348; Emp24_gp25L_D24.
CC InterPro; IPR009038; GOLD.
CC Pfam; PF006050; CRAL_TRIO; 1.
CC Pfam; PF03765; CRAL_TRIO_N; 1.
CC Pfam; PF01105; EMP24_GP25L; 1.
CC PRINTS; PR00180; CRETINALDHP.
CC SMART; SM00516; SEC14; 1.
CC PROSITE; PS50191; CRAL_TRIO; 1.
CC PROSITE; PS50866; GOLD; 1.
CC 3D-structure; Activator; Alternative splicing; Lipid-binding;
KW Nuclear protein; Transcription; Transcription regulation; Transport.
FT DOMAIN 76 249 CRAL-TRIO.
FT DOMAIN 275 383 Missing (in isoform 2).
FT VARSPLIC 361 403 /FTId=VSP_006031.
FT CONFLICT 36 36 Y -> H (in Ref. 2).
FT SEQUENCE 403 AA; 46145 MW; D846747EC8D1513E CRC64;
SQ
Query Match 90.6%; Score 1827.5; DB 1; Length 403;
Best Local Similarity 87.1%; Pred No. 2,7e-140;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;
QY 1 MSGRGVGLSPRQKEALAK-----PEASTCRSRPCSGMWSSES-----39
Db 1 MSGRGVGLSPRQKEALAKFRENVDVLPALPNPDDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTTSLAWOPPEVIQOYLSGGMGVDLDCGFWYDIIGPKDAKGLLPSASKQDL 95
Db 61 EFRKQIDNTIISWQPPEVIQOYLSGGMGVDLDCGFWYDIIGPKDAKGLLPSASKQDL 120
QY 96 LRTKMRCELLQECACHQTTKLGRKVEITITTYDCEGLGKLHLKWPVAVEAYGFLCMFEE 155
Db 121 LRTKMRCELLQECACHQTTKLGRKVEITITTYDCEGLGKLHLKWPVAVEAYGFLCMFEE 180
QY 156 NYPETLKLFFVVKAPKLPFPVYVNLKPFSEDRKKIMVGLGANYKEVLLKHISPOQVPVE 215
Db 181 NYPETLKLFFVVKAPKLPFPVYVNLKPFSEDRKKIMVGLGANYKEVLLKHISPOQVPVE 240

```

RC TTSUUP=Liver;
RX MEDLINE=42388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]

RC	TISSUE=Liver:		
RA	Director MGC Project:		
RL	Submitted (FEB-2005) to the EMBL/GenBank/DBDJ databases:		
DR	EMBL; BC089785; AAH9785.1; ; MRNA.		
DR	InterPro; IPR001251; CRAL_TRIO_C.		
DR	InterPro; IPR008273; CRAL_TRIO_N.		
DR	InterPro; IPR000348; Emp24_gp25L_p24.		
DR	InterPro; IPR000938; GOLD.		
DR	InterPro; IPR001071; RetBind/tocTrans.		
DR	Pfam; PF00650; CRAL_TRIO; 1.		
DR	Pfam; PF03765; CRAL_TRIO_N; 1.		
DR	Pfam; PF01105; EMP24_GP25L; 1.		
DR	PRINTS; P000180; CRETINALDHP.		
DR	SMART; SM00516; SEC14; 1.		
DR	PROSITE; PS50191; CRAL_TRIO; 1.		
DR	PROSITE; PS50866; GOLD; 1.		
SQ	SEQUENCE 403 AA, 46205 MW; D83CCF10DBA4A4837 CRC64;		
		Query Match	85.6%; Score 1726.5; DB 2; Length
		Best Local Similarity	81.1%; Pred. No. 4.7e-132;
		Matches 327; Conservative 25; Mismatches 26; Indels	
QY	1 MSGRVCDSLSPKKEALAK-----PEASTCRRSPCCS		

```
Qy      1 MSGRVDLSPQXEALAK-----PEASTCRSRPPCSGSMWSSS----- 39
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
RC STRAIN=Wistar; TISSUE=Liver;
RX PubMed=1126224; DOI=10.1073/pnas.041620398;
RA Shibata N., Arita M., Misaki Y., Dohmae N., Takio K., Ono T.,
RA Inoue K., Arai H.;
RT "Supernatant protein factor, which stimulates the conversion of
RT squalene to lanosterol, is a cytosolic squalene transfer protein and
RT enhances cholesterol biosynthesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:2244-2249(2001).
CC -!- FUNCTION: Carrier protein. Binds to some hydrophobic molecules and
CC promotes their transfer between the different cellular sites.
CC Binds with high affinity to alpha-tocopherol. Also binds with a
CC weaker affinity to other tocopherols and to tocotrienols. May have
CC a transcriptional activatory activity via its association with
CC alpha-tocopherol (By similarity). Probably recognizes and binds
CC some squalene structure, suggesting that it may regulate
CC cholesterol biosynthesis by increasing the transfer of squalene to
CC a metabolic active pool in the cell.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic in absence of alpha-tocopherol,
CC and nuclear in presence of alpha-tocopherol (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q99MS0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q99MS0-2; Sequence=Not described;
CC Name=3;
CC IsoId=Q99MS0-3; Sequence=Not described;
CC -!- TISSUE SPECIFICITY: Widely expressed. High expression in liver and
CC small intestine.
CC -!- SIMILARITY: Contains 1 CRAL-TRIO domain.
CC -!- SIMILARITY: Contains 1 GOLD domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF309558; AAK16405.1; -; mRNA.
CC SMR; Q99MS0; 1-396.
CC Ensembl; ENSRNOG00000004672; Rattus norvegicus.
CC DR; GGD; 621779; Sec14L2.
CC DR; GO; GO:0005829; C:cytosol; IDA.
CC DR; GO; GO:0005634; C:nucleus; ISS.
CC DR; GO; GO:0008047; F:enzyme activator activity; IDA.
CC DR; GO; GO:0005543; F:phospholipid binding; ISS.
CC DR; GO; GO:0008320; F:protein carrier activity; ISS.
CC DR; GO; GO:0016563; F:transcriptional activator activity; ISS.
CC DR; GO; GO:0008431; F:vitamin E binding; ISS.
CC DR; GO; GO:0045542; P:positive regulation of cholesterol biosynth. . .; IPI.
CC DR; GO; GO:0045893; P:positive regulation of transcription, DNA-d. . .; ISS.
CC DR; InterPro; IPR001071; CRAL_bd toc tran.
CC DR; InterPro; IPR001251; CRAL_bd TRIO_C.
CC DR; InterPro; IPR008273; CRAL_bd TRIO_N.
CC DR; InterPro; IPR00348; Emp24_gp25L_p24.
CC DR; InterPro; IPR009038; GOLD.
CC DR; Pfam; PF03765; CRAL_TRIO; 1.
CC DR; Pfam; PF01105; EMP24_GP25L; 1.
CC DR; PRINTS; PR00180; CRETINALDHP.
CC DR; SMART; SM00516; SEC14; 1.
CC DR; PROSITE; PS50191; CRAL_TRIO; 1.
CC DR; PROSITE; PS50866; GOLD; 1.
CC Activator; Alternative splicing; Direct protein sequencing;
CC Lipid-binding; Nuclear protein; Transcription;
CC Transcription regulation; Transpos.
CC DOMAIN 76 249
CC FT DOMAIN 275 383
CC SEQUENCE 403 AA; 46166 MW; 413CC8BC4E3A45BD CRC64;
CC SQ

RESULT 6
S14L2 RAT
ID S14L2 RAT STANDARD; PRT; 403 AA.
AC Q99MS0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SEC14-like protein 2 (Alpha-tocopherol associated protein) (TAP)
DE (Supernatant protein factor) (SPF) (Squalene transfer protein).
DE Name=Sec14L2;
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10116;
RN (1)
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 19-51 AND 231-252.

Query Match 85.5%; Score 1724.5; DB 2; Length 403;
Best Local Similarity 80.9%; Pred. No. 6.8e-132; Indels 25; Gaps 2;
Matches 326; Conservative 27; Mismatches 25;

QY 1 MSGRVGDLSPQKEALAK-----PEASTCRSRPCSGSMWSSS----- 39
DB 1 MSGRVGDLSPQKEALAKFRENVDVLPPLPNDPDYLLRLWLRARSFDLQSEAMLRKHV 60
QY 40 ----KRTLTTSLAWQPPEVITQVLSGCMGYDLDGCPVWYDIIGPKDAKGLLFSASKQDL 95
DB 61 EFRQKQDKIISQNPPEVITQVLSGCMGYDLDGCPVWYDIIGPKDAKGLLFSASKQDL 120
QY 96 LRTKMRCELLQECANHTTKGKGVETITIVDCEGLGKHLWKPAVEAYGFLCMFEE 155
DB 121 LRTKMRCELLQECIQOTTKLGKGIETITIVDCEGLGKHLWKPAVEAYGFLTMFEE 180
QY 156 NYPETLKLRFVVKAPKLPVAYNLIKPFLSEDTTRKIMVLGANYKEYLLKHISPDQVPE 215
DB 181 NYPETLKLRFVVKAPKLPVAYNLIKPFLSEDTTRKIMVLGANYKEYLLKHISPDQVPE 240
QY 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVYRDQVKQYHSVQISRGSSHQVEYIILFPG 275
DB 241 YGGTMTDPDGNPKCKSKINYGDDIPKQYVYRDQVKQYHSVQISRGSSHQVEYIILFPG 300
QY 276 CVLRQFMSDGDVGFGLFKTKMGERQAGEMTEVLPNORYNSHLVPEDGTLTCSDPGI 335
DB 301 CVLRQFMSDGDVGFGLFKTKMGERQAGEMTEVLPNORYNSHLVPEDGTLTCSDPGI 360
QY 336 YVLRFDNTYSPIHAKVNFTEVLLLPKASEKMKQLGAGTPK 378
DB 361 YVLRFDNTYSPIHAKVNFTEVLLLPKASEKMKQLGAGTPK 403

RESULT 6
S14L2 RAT
ID S14L2 RAT STANDARD; PRT; 403 AA.
AC Q99MS0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SEC14-like protein 2 (Alpha-tocopherol associated protein) (TAP)
DE (Supernatant protein factor) (SPF) (Squalene transfer protein).
DE Name=Sec14L2;
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10116;
RN (1)
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 19-51 AND 231-252.
```

```
Query Match      85.4%; Score 1722.5; DB 1; Length 403;
Best Local Similarity 81.1%; Pred. No. 9.9e-132;
Matches 327; Conservative 25; Mismatches 26; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPKQKEALAK-----PEASTCRSRPPCGSGMWSES----- 39
DB 1 MSGRVGDLSPKQKEALAKFRENVDVLPALPNPDDYFLRLWLRLARSFDLQKSEAMLRKHV 60

QY 40 ----KRTLTTSIAWQPPEVIQOYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
DB 61 EFRKQKIDIKIISWQPPEVIQOYLSGGRCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120

QY 96 LRTKRECELLLOECACHQTTKLRKVETITIIYDCGELGLKHLWKPAVEAYGEFLCMFEE 155
DB 121 LRTKMRDCELLLOECQTTAKLGKKTETITIIYDCGELGLKHLWKPAVEAYGEFTLMFEE 180

QY 156 NYPETLKRFLFVVKAPKLPFVAYNLKLPFSEDTTRKKIMVLGANYKEVLLKHISPDQVPVE 215
DB 181 NYPETLKRFLFVVKAPKLPFVAYNLKLPFSEDTTRKKIMVLGANYKEVLLKHISPDQVPVE 240

QY 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVYRDQVQKQYEHVSQVSRGSSHQVEYELPFG 275
DB 241 YGGTMTDPDGNPKCKSKINYGDDIPKQYVYRDQVQKQYEHVSQVSRGSSHQVEYELPFG 300

QY 276 CVLRWQFMSGDADVGFIFLTKMGERQAGEMTEVLPNQRVNSHLVPEDGTLTCSDPGI 335
DB 301 CVLRWQFMSGSDVGFIFLTKMGERQAGEMTEVLPNQRVNSHLVPEDGTLTCSDPGI 360

QY 336 YVLRFDNTYSFIHAKVNFTEVLLPDKASEEKMQLGAGTPK 378
DB 361 YVLRFDNTYSFIHAKVNFTEVLLPDKAAEKLNQOGAVTPK 403

RESULT 7
Q867A0_BOVIN
ID Q867A0_BOVIN PRELIMINARY; PRT; 403 AA.
AC Q867A0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Liver tocopherol-associated protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20400504; PubMed=10829015; DOI=10.1074/jbc.M000851200;
RA Zimmer S., Stocker A., Sarbolouki M.N., Spycher S.E., Sassoon J.,
RA Azzi A.;
RT "A novel human tocopherol-associated protein: cloning, in vitro
RT expression, and characterization.";
RL J. Biol. Chem. 275:25672-25680(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Breyer I., Schittny J.C., Schuerch S., Zimmermann A., Stocker A.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432353; AAC031942.1; -; mRNA.
DR HSSP; 076054; 106U.
DR SMR; Q867A0; 1-397.
DR GO; GO:0003622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0008320; F:protein carrier activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008886; P:intracellular protein transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001251; CRAL TRIO C.
DR InterPro; IPR008273; CRAL TRIO N.
DR InterPro; IPR000348; Emp24_gp25L_p24.
DR InterPro; IPR009038; GOLD.
DR InterPro; IPR001071; RetBind/tocTrans.
```

```
DR Pfam; PF00650; CRAL_TRIO; 1.
DR Pfam; PF03765; CRAL_TRIO_N; 1.
DR Pfam; PF01105; EMP24_GP25L; 1.
DR PRINTS; PR00180; CRETINALDHP.
DR SMART; SM00516; SEC14; 1.
DR PROSITE; PS00191; CRAL_TRIO; 1.
DR PROSITE; PS00866; GOLD; 1.
SQ SEQUENCE 403 AA; 46200 MW; 67C28BFC173E1CD9 CRC64;

Query Match      83.5%; Score 1685.5; DB 2; Length 403;
Best Local Similarity 78.2%; Pred. No. 1e-128;
Matches 315; Conservative 34; Mismatches 29; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPKQKEALAK-----PEASTCRSRPPCGSGMWSES----- 39
DB 1 MSGRVGDLSPKQKEALAKFRENVDVLPALPNPDDYFLRLWLRLARNFNLQKSEAMLRKHV 60

QY 40 ----KRTLTTSIAWQPPEVIQOYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
DB 61 EFRKQKIDINMSWQPPEVQOYLSGGMCGYDLGSPDWYDIIGPLDAKGLLFSASKQDL 120

QY 96 LRTKRECELLLOECACHQTTKLRKVETITIIYDCGELGLKHLWKPAVEAYGEFLCMFEE 155
DB 121 FKTMRDCELLLOECVROTETKMGKKEATTIIYDCGELGLKHLWKPAVEAYGEFLCMFEE 180

QY 156 NYPETLKRFLFVVKAPKLPFVAYNLKLPFSEDTTRKKIMVLGANYKEVLLKHISPDQVPVE 215
DB 181 NYPETLKRFLFVVKAPKLPFVAYNLKLPFSEDTTRKKIQVLGNANWKEVLLKYISPDQLPVE 240

QY 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVYRDQVQKQYEHVSQVSRGSSHQVEYELPFG 275
DB 241 YGGTMTDPDGNPKCKSKINYGDDIPKQYVYRDQVQKQYEHVSQVSRGSSHQVEYELPFG 300

QY 276 CVLRWQFMSGDADVGFIFLTKMGERQAGEMTEVLPNQRVNSHLVPEDGTLTCSDPGI 335
DB 301 CVLRWQFMSGSDIGFIFLTKMGERQAGEMREVLPQRYNAHLVPEDGSLTCSDPGI 360

QY 336 YVLRFDNTYSFIHAKVNFTEVLLPDKASEEKMQLGAGTPK 378
DB 361 YVLRFDNTYSFIHAKVNFTEVLLPDKALEEKMQLGAVTPK 403

RESULT 8
S14L2_BOVIN
ID S14L2_BOVIN STANDARD; PRT; 387 AA.
AC P58875;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SEC14-like protein 2 (Alpha-tocopherol associated protein) (TAP)
DE (bTAP) (Fragment).
GN Name=SEC14L2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Meadus J., MacInnis R., Dubeski P., Hidioglou N., Madere R.;
RT "Induction of hepatic tocopherol associated protein (TAP) mRNA but not
RT alpha-tocopherol transfer protein (TTP) mRNA in cattle fed increasing
RT levels of vitamin E.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PROTEIN SEQUENCE OF 178-195 AND 335-353.
RX MEDLINE=20400504; PubMed=10829015; DOI=10.1074/jbc.M000851200;
RA Zimmer S., Stocker A., Sarbolouki M.N., Spycher S.E., Sassoon J.,
RA Azzi A.;
RT "A novel human tocopherol-associated protein: cloning, in vitro
RT expression, and characterization.";
RL J. Biol. Chem. 275:25672-25680(2000).
```

CC -1- FUNCTION: Carrier protein. Binds to some hydrophobic molecules and
 CC promotes their transfer between the different cellular sites.
 CC Binds with high affinity to alpha-tocopherol. Also binds with a
 CC weaker affinity to other tocopherols and to tocotrienols. May have
 CC a transcriptional activator activity via its association with
 CC alpha-tocopherol. Probably recognizes and binds some squalene
 CC structure, suggesting that it may regulate cholesterol
 CC biosynthesis by increasing the transfer of squalene to a metabolic
 CC active pool in the cell (By similarity).
 CC
 CC -1- SUBUNIT: Monomer.
 CC
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic in absence of alpha-tocopherol,
 CC and nuclear in presence of alpha-tocopherol (By similarity).
 CC
 CC -1- SIMILARITY: Contains 1 CRAL-TRIO domain.
 CC
 CC -1- SIMILARITY: Contains 1 GOLD domain.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC -----
 CC ENBL; AF487977; AAL90886.1; -; mRNA.
 CC SMR; P58875; 1-387.
 CC GO; GO:0005737; C:cytoplasm; ISS.
 CC GO; GO:0005634; C:nucleus; ISS.
 CC GO; GO:0005543; F:phospholipid binding; ISS.
 CC GO; GO:0008320; F:protein carrier activity; ISS.
 CC GO; GO:0016563; F:transcriptional activator activity; ISS.
 CC GO; GO:0008431; F:vitamin E binding; ISS.
 CC GO; GO:0045893; P:positive regulation of transcription, DNA-d. .; ISS.
 CC GO; GO:0045540; P:regulation of cholesterol biosynthesis; ISS.
 CC InterPro; IPR001071; CRAL_bd toc tran.
 CC InterPro; IPR001251; CRAL_bd TRIO C.
 CC InterPro; IPR008273; CRAL_bd TRIO N.
 CC InterPro; IPR000348; Emp24_GP25L_P24.
 CC InterPro; IPR009038; GOLD.
 CC Pfam; PF00650; CRAL_TRIO; 1.
 CC Pfam; PF03785; CRAL_TRIO_N; 1.
 CC Pfam; PF01105; EMP24_GP25L; 1.
 CC PRINTS; PR00180; CRETINALDHP.
 CC SMART; SM00516; SEC14; 1.
 CC PROSITE; PS50191; CRAL_TRIO; 1.
 CC PROSITE; PS50866; GOLD; 1.
 CC Activator; Direct protein sequencing; Lipid-binding; Nuclear protein;
 CC Transcription; Transcription regulation; Transport.
 CC
 CC FT DOMAIN 76 249
 CC FT DOMAIN 275 383
 CC FT CONFLICT 193 193 K -> G (in Ref. 2).
 CC FT CONFLICT 343 343 S -> N (in Ref. 2).
 CC FT NON_TER 387 387
 CC SEQUENCE 387 AA; 44406 MW; FAE72E8A83CB9E9C CRC64;
 CC
 CC Query Match 80.1%; Score 1616.5; DB 1; Length 387;
 CC Best Local Similarity 77.5%; Pred. No. 4.1e-123;
 CC Matches 300; Conservative 35; Mismatches 27; Indels 25; Gaps 2;
 CC
 CC QY 1 MSGRVGDLSPKQKALAK-----PEASTCSRPRPCSGSMWSSS----- 39
 CC DB 1 MSGRVGDLSPKQKALAKFRENQDVLFPALPNPDYFLRLWRARNFNLOKSEAMLRKHV 60
 CC QY 40 ---KRTLTTSIAWQPPVIOQYLSGGMCGYDLGCPWVDYIGPKDAKGLLPFSASKQDL 95
 CC DB 61 EFRKQKIDNIMSWQPPVEVQYLSGGMCGYDLGSGPIWYDIIGPLDAKGLLLSASKQDL 120
 CC QY 96 LRTKMRCELLQSCAHQTTKLGKRVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
 CC DB 121 FTKMRCELLQECVQTRQKWKKKIATIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
 CC QY 156 NYPETLKRFLFVVKAPKLPFVAYNLKIPPLSDTRKKIMVLGANYKEVLLKHSIDQVPVE 215
 CC DB 181 NYPETLKRFLFVVKAPKLPFVAYNLVKKPLSDTRKKIQVLGANYKEVLLKYSIDQLPEV 240
 CC QY 216 YGTTWTDGPNPKCKSKINYGDIIPKYYVRDQVQYEHVSQISRGSSHQVEYELLFPG 275

DB 241 YGTTWTDGPNPKCKSKINYGDIIPKYYVRDQVQYEHVSQISRGSSHQVEYELLFPG 300
 QY 276 CVLRQGFMSDGDADVGFGLFKTKGERORAGRMTEVLNQRVNSHLVPDGTLTCSDDPGI 335
 DB 301 CVLRQGFMSDGDSDIGFGFLTKTKVGERQAGEREVLFSQRIASHLVDPGSGISLTCSDDPGI 360
 QY 336 YVLRFDNTYSFTHAKKKNFTVEVLLPD 362
 DB 361 YVLRFDNTYSFTHAKKKNFTVEVLLPD 387
 RESULT 9
 Q6PD61 HUMAN
 ID Q6PD61 HUMAN PRELIMINARY; PRT; 392 AA.
 AC Q6PD61
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE SEC14L2 protein.
 GN Name=SEC14L2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; and
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RC Director MGC Project;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC058915; AAH58915.1; -; mRNA.
 DR SMR; Q6PD61; 1-360
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR GO; GO:0008320; F:protein carrier activity; IEA.
 DR GO; GO:0019841; F:retinol binding; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006886; P:intracellular protein transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR GO; GO:0007601; P:visual perception; IEA.
 DR InterPro; IPR001251; CRAL_TRIO C.
 DR InterPro; IPR008273; CRAL_TRIO N.
 DR InterPro; IPR000348; Emp24_GP25L_P24.
 DR InterPro; IPR001071; RetBind/tocTrans.
 DR Pfam; PF00650; CRAL_TRIO; 1.
 DR Pfam; PF03785; CRAL_TRIO_N; 1.
 DR Pfam; PF01105; EMP24_GP25L; 1.

Query Match	78.3%	Score 1580.5	DB 2	Length 392
Best Local Similarity	83.94	Pred. No. 3.6e-120		
Matches 302	Conservative 13	Mismatches 20	Indels 25	Gaps 2
<p>DR PRINTS; PR00180; CRETINALDHPB.</p> <p>DR SMART; SM00516; SEC14; 1.</p> <p>DR PROSITE; PS50191; CRAL/TRIO; 1.</p> <p>KW Hypothetical protein.</p> <p>SQ SEQUENCE 392 AA; BDC2D956378ABE5 CRC64;</p>				
QY	1	MSGRVGDLSPKQKALAK-----PEASTCRSRPCSGSMWSSS-----	39	
DB	1	MSGRVGDLSPKQKALAKFENVQDVLPALPNDDYFLLRWLRARSFDLQKSEAMLRKHV	60	
QY	40	----KRTLTSLAWQPPEVIQQYLSGMCYDLDGCPVWYDIIGPKDAKGLLFSASKODL	95	
DB	61	EPRKQKIDINISWQPPEVIQQYLSGMCYDLDGCPVWYDIIGPLDAKGLLFSASKODL	120	
QY	96	LRTKMRCELLLOECAHQTTLKGRKVEITIIIVDCEGLGKHLWKPAVEAYGEFLCMFEE	155	
DB	121	LRTKMRCELLLOECAHQTTLKGRKVEITIIIVDCEGLGKHLWKPAVEAYGEFLCMFEE	180	
QY	156	NYPETLKRLLFWKAPKLPVAYNLIKPFSLSEDRKKIMVLGANYKEVLLKHISPDQVPE	215	
DB	181	NYPETLKRLLFWKAPKLPVAYNLIKPFSLSEDRKKIMVLGANYKEVLLKHISPDQVPE	240	
QY	216	YGGTMTDPDGNPKCKSKINYGDIIPKRYVYRDQVKQYHSHVQISRGSSHQVEYIILPPG	275	
DB	241	YGGTMTDPDGNPKCKSKINYGDIIPKRYVYRDQVKQYHSHVQISRGSSHQVEYIILPPG	300	
QY	276	CVLRWQFMSDQADVGFQIFLTKMGERQORAGEMTEVLPNQRYNSHLVPEDGTLTCSDDGI	335	
DB	301	CVLRWQFMSDQADVGFQIFLTKMEERQORAGEMTEVLPNQRYNSHLVPEDGTLTCSDDGI	360	
<p>RESULT 11</p> <p>SI4L3_HUMAN STANDARD; PRT; 400 AA.</p> <p>AC Q3UDX4;</p> <p>DT 10-OCT-2003 (Rel. 42, Created)</p> <p>DT 10-OCT-2003 (Rel. 42, Last sequence update)</p> <p>DT 10-MAY-2005 (Rel. 47, Last annotation update)</p> <p>DE SEC14-like protein 3 (Tocopherol-associated protein 2).</p> <p>GN Name=SEC14L3; Synonyms=TA2P;</p> <p>OS Homo sapiens (Human).</p> <p>OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;</p> <p>OC Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini; Homnidae;</p> <p>OC Homo.</p> <p>OX NCBI_TaxID=9606;</p> <p>RN [1]</p> <p>RP NUCLEOTIDE SEQUENCE.</p> <p>RX MEDLINE=22642877; PubMed=12757856; DOI=10.1016/S0891-5849(03)00173-4;</p> <p>RA Kempna P., Zingg J.-M., Ricciarelli R., Hierl M., Saxena S., Azzi A.;</p> <p>RT "Cloning of novel human SEC14p-like proteins: ligand binding and functional properties";</p> <p>RL Free Radic. Biol. Med. 34:1458-1472 (2003).</p> <p>[2]</p> <p>RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].</p> <p>RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;</p> <p>RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,</p> <p>RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,</p> <p>RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beales O.P.,</p> <p>RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,</p> <p>RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,</p> <p>RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,</p> <p>RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,</p> <p>RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,</p> <p>RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,</p> <p>RA Gilbert J.G.R., Goward M.E., Grahame D.V., Griffiths M.N.D., Hall C.,</p> <p>RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,</p> <p>RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,</p> <p>RA Laird G.K., Langford C.F., Leverisa M.A., Lloyd C., Lloyd D.M.,</p> <p>RA Martyn I.D., Madsregheh-Mohammadi M., Matthews L.H., Mccann O.T.,</p>				


```
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey K., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Swann R.M.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin L., Walli M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuura S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kanton S., Lai H., Lao H.T.,
RA Lewis J., Lewis S., Lin S.P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren O., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wansley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Serousi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilihan Y., Wright H.,
RA Nature 402:489-495(1999).
CC -!- FUNCTION: Probable hydrophobic ligand-binding protein; may play a
CC role in the transport of hydrophobic ligands like tocopherol,
CC squalene and phospholipids.
CC -!- SIMILARITY: Contains 1 CRAL-TRIO domain.
CC -!- SIMILARITY: Contains 1 GOLD domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AY158086; AAC21870.1; -; mRNA.
CC ENBL; AC004832; AAF19258.1; -; Genomic_DNA.
CC SMK; Q9UDX4; 1-397.
CC DR Ensembl; ENSG00000100012; Homo sapiens.
CC DR HGNC; HGNC:18655; SEC14L3.
CC DR InterPro; IPR001071; CRAL_bd_toc_tran.
CC DR InterPro; IPR001251; CRAL_bd_TRIO_C.
CC DR InterPro; IPR008273; CRAL_bd_TRIO_N.
CC DR InterPro; IPR000348; Emp24_gp25L_p24.
CC DR InterPro; IPR009038; GOLD.
CC DR Pfam; PF00650; CRAL_TRIO; 1.
CC DR Pfam; PF03765; CRAL_TRIO_N; 1.
CC DR Pfam; PF01105; EMP24_gp25L; 1.
CC DR PRINTS; PR00180; CRETINALDHP.
CC DR PROSITE; PSS0191; CRAL_TRIO; 1.
CC DR PROSITE; PSS0866; GOLD; 1.
CC DR Lipid-binding; Polymorphism; Transport.
CC KW DOMAIN 76 249 CRAL-TRIO.
CC FT DOMAIN 275 383 GOLD.
CC FT VARIANT 335 335 D->E (in dbSNP:2240345).
CC FT /FTID=VAR_022097.
CC SEQUENCE 400 AA; 46048 MW; 07F880D25B66CC19 CRC64;
Query Match 70.8%; Score 1428; DB 1; Length 400;
Best Local Similarity 66.4%; Pred No. 9.8e-108;
Matches 267; Conservative 51; Mismatches 56; Indels 28; Gaps 4;
QY 1 MSGRVGDLSPRQKEALAK-----PEASTCSRPRPCSGMWSSES----- 39
Db 1 MSGRVGDLSPRQKEALAKFRENVDVLPALPNPDYFLLRWLARNFDLQKSEALLRKYM 60
QY 40 --KRTLTS--LAWQPEVITQQYLSGNGCGYDLDCGPVWYDIIGPKDAKGLLFSASKQDL 95
```

```
Db 61 EPRKMTIDHILDWQPEVIQKMPGGLCGYDRDCGPVWYDIIGPKGLLFSVTQDL 120
QY 96 LETKRECELLQECACHOTTKLGRKVEITIIYDCGGLGKHLKWPAAVEAYGEFLCMPRE 155
Db 121 LKTRNRDCERILHECDLQTERLGKKIETIYVLPFCGGLGKHLKWPAAVEAYGEFLLEE 180
QY 156 NYPETLKRFLVVKAPKLPFVAYNLKLPFLSEDRKKIMVLGANYKEVLLKHSPOQVPE 215
Db 181 NYPETLKRFLVVKAPKLPFVAYNLKLPFLSEDRKKIMVLGANYKEVLLKHSPELPAQ 240
QY 216 YGGTWTDDGPNPKCKSKINYGDIIPKYYVRDQVQOYHSHVQISRGSSHOVEYELFPG 275
Db 241 FGGTLTDPGPNPKCLTKINYGGEIPKSMYVRDQVQOYHSHVQISRGSSHOVEYELFPG 300
QY 276 CVLRWQFSDGADVGGIFLTKMGORAGMTVLNQRVNSHLVPBDGTLTCSDDPGI 335
Db 301 CVLRWQFSDGADIGFGLTKMGORAGMTDVLFSQRNNAHVDPDGNLTCEAGV 360
QY 336 YVLRFDNTYSFTHAKKVNFTVEVLLPDKASEEMKQLGAGTP 377
Db 361 YVLRFDNTYSFVHAKKVSFTVEVLLPDEGMQKYDEL---TP 399
RESULT 12
O6ISB2 HUMAN PRELIMINARY; PRT; 400 AA.
AC O6ISB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SEC14-like 3.
GN Name=SEC14L3;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN (1) _TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX MEDLINE=23288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC069641; AAH69641.1; -; mRNA.
DR HSSP; P24280; IAUa.
DR SMR; Q6ISB2; 1-397.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008320; F:protein carrier activity; IEA.
```

DR	GO; GO:0005215; F:transporter activity; IEA.
DR	GO; GO:0006886; P:intracellular protein transport; IEA.
DR	GO; GO:0006910; P:transport; IEA.
DR	InterPro; IPRO01071; CRAL_bd_toc_tran.
DR	InterPro; IPRO01251; CRAL_bd_TRIO_C.
DR	InterPro; IPRO08273; CRAL_bd_TRIO_N.
DR	InterPro; IPRO03048; Emp21_gp25L_P24.
DR	InterPro; IPRO09038; GOLD.
DR	Pfam; PF00650; CRAL_TRIO_1.
DR	Pfam; PF03765; CRAL_TRIO_N; 1.
DR	Pfam; PF01105; EMP24_GP25L; 1.
DR	PRINTS; PR00180; CRETINALDHP.
DR	SMART; SM00516; SEC14; 1.
DR	PROSITE; PSS0191; CRAL_TRIO_1.
DR	PROSITE; PS50866; GOLD; 1.
SQ	SEQUENCE 400 AA; 46080 MW; 8FBEC87350F53C10 CRC64;

Query Match		70.7%; Score 1426; DB 2; Length 400;
Best Local Similarity		66.4%; Pred. No. 1.4e-107; Indels 28; Gaps 4;
Matches 267; Conservative 50; Mismatches 57;		

Qy	1	MSGRVGDLSPRQKEALAK-----PEASTCRSRRCSCGSMWSES-----	39
Dd	1	MSGRVGDLSPKQAETLAKFRENVQDVLPALPNPDYFLRLWLARNPDLQKSETLLRKYM	60
Qy	40	--KRTLTS--LAWQPBEVITQQYLSGCMGVLDGCPVWYDIIGPKDAKGLLFPASKQDL	95
Dd	61	EFRKTMDIHILWDQPPEVIQKNYPGGLCGYDRDGCPVWYDITGFLDPKGLLFSVTQKL	120
Qy	96	LRTKWRECELLQRCAHQTTLGRKRVETIIYYCEGLGLKHLMKPAEAYGEFLCMPEE	155
Dd	121	LKTWRDCERILHECDIQTERLGKKIETIVMIFCEGLGLKHFVKPLVEVYQEFGLLEE	180
Qy	156	NYPETLKRLFVVKAPKLPFAYNLIKPFLSEDTRRKKIMVLGANYSKVELLKHSIDQPVPE	215
Dd	181	NYPETLKFMLIVKATKLPFGVGNLMKPFLSDTRRKIIVLGNMWKEGLLKLISPEELPAQ	240
Qy	216	YGWTMTDPDGNPKCKSKINYGDDIPRKYVVVDQVKQYEHSVOITSRGSSHQVEHILPPG	275
Dd	241	PGFTLTDPDGNPKCLTKINYGGEIPKSNMYVRDQVKTYEHVSQINRGSSHQVEHILPPG	300
Qy	276	CVLRWQFMSDCADVGFGLFKTKNGERORAGEMTEVLNQRYNSHLVPEDGTGLTCSDPGI	335
Dd	301	CVLRAWQSDDADTGFGVFLTKKGQRORAGEMTEVLPQRYNAHWVEDGNLTCSAGV	360
Qy	336	YVLRFDNTSYTFIAKKNVFTVEVLLPDKASBEKMQLGAGTP	377
Dd	361	YVLRFDNTSYTFVAHKYSFTVEVLLPDEGMOKYDKEL---TF	399

RESULT 13
Q6XC17 HUMAN
ID Q6XC17 HUMAN PRELIMINARY; PRT; 400 AA.
AC Q6XC17;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SEC14-like protein 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RP PubMed=15040456; DOI=10.1023/B:MOLE.0000013504.88003.32;
RA Ye X., Ji C., Yin G., Tang R., Zeng L., Gu S., Ying K., Xie Y.,
RA Zhao R.C., Mao Y.;
RT "Characterization of a human Sec14-like protein cDNA SEC14L3 highly
homologous to human SPF/TAP.";
RL Mol. Biol. Rep. 31:59-63 (2004).
DR EMBL; AY240872; AAO52677.1; -, mRNA.
DR HSSP; P24280; 1AUA.

DR	SMR; Q6XC17; 1-397.
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0008289; F:lipid binding; IEA.
DR	GO; GO:0008320; F:protein carrier activity; IEA.
DR	GO; GO:0005215; F:transporter activity; IEA.
DR	GO; GO:0006886; P:intracellular protein transport; IEA.
DR	GO; GO:0006810; P:transport; IEA.
DR	InterPro; IPR001251; CRAL_TRIO_C.
DR	InterPro; IPR008273; CRAL_TRIO_N.
DR	InterPro; IPR000348; Emp24_gp25L_p24.
DR	InterPro; IPR009038; GOLD
DR	InterPro; IPR001071; RetBind/tocTrans.
DR	Pfam; PF00650; CRAL_TRIO; 1.
DR	Pfam; PF03765; CRAL_TRIO_N; 1.
DR	Pfam; PF01105; EMP24_GP25L; 1.
DR	PRINTS; PR00180; CRETINALDHP.
DR	SMART; SM00516; SEC14; 1.
DR	PROSITE; PSS0191; CRAL_TRIO; 1.
DR	PROSITE; PS50866; GOLD; 1.
SQ	SEQUENCE 400 AA; 46084 MW; 615AA9D931FC135F CRC64;
 Query Match 70.6%; Score 1425; DB 2; Length 400;	
Best Local Similarity 66.4%; Pred No. 1.7e-107;	
Matches 267; Conservative 50; Mismatches 57; Indels 28; Gaps 4	
Qy	1 MSGRVGDSLPRQKEALAK-----PEASTCRRSPCGSGMMSSES----- 39 : : :
Dd	1 MSGRVGDSLPKQAETLAKPRENVQDVLPALPNPDYFLRLWLARNFDLQKSEALLRKYM 60 : : :
Qy	40 --KRTLTS--LAWQPPEVTQQYLSGMGYDDLDCPPWVDIITGPKDAKGLLFSASKODL 95 :: : : : : : : : :
Dd	61 EFRKTMDIHLDWQPPEVIQKYMPGGLCGYDRDGCPVWYDITGPDPKGLLFSVTKODL 120 : : : : : : : : :
Qy	96 LRTKMRECELLQSCAHOTTKLGSRKVFTITIYDCEGLGHLWKPAVEAYGEFLCMPEE 155 : : : : : : : : :
Dd	121 LKTKMRDCEKLILHECDLQTERLGKKITIVMIFCEGLGLGHFKWPLVEVYQEFPGLLEE 180 : : : : : : : : :
Qy	156 NYPETLKRLFVKAPKLPVAYNLIKPFLSEDTRKKIMVLGANKEYVELLKHSIDPQVPVE 215 : : : : : : : : :
Dd	181 NYPETLKFMLIVKATKLPFGYNLMKPFLESDTRRKIIVLGNMKEGLLKLISPEELPAQ 240 : : : : : : : : :
Qy	216 YGGTMTDPDGNPKCKSKINYGDDIPRYKYVRDQVKQYEHSVOISRSSHQVEVEILPG 275 : : : : : : : : :
Dd	241 FGGLTPDGNPKCLTKINYGGEIPKMYVRDQVKTYEHVSQVINRGSSHQVEVEILPG 300 : : : : : : : : :
Qy	276 CVLRQWFMSDGDGVGFGLTKMGERORAGEMTEVLPNQRYNSHLYVPEDGTLTCSDPGI 335 : : : : : : : : :
Dd	301 CVLRQWFSSDGDADTFGFFLTKMGERORAGEMTEVLPNQRYNAHMVEDGNLTCSEAGV 360 : : : : : : : : :
Qy	336 YVLRFDNTYSFIHAKNYFTVEVLLPKASEEKMKQLGAGTF 377 : : : : : : : : :
Dd	361 YVLRFDNTYSFVHAKNYSFTVEVLLPDGWMQYDKEL---TP 399 : : : : : : : : :
 RESULT 14	
ID	S14L3 RAT
AC	Q9Z1J8; STANDARD; PRT; 400 AA.
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	10-MAY-2005 (Rel. 47, Last annotation update)
DE	SEC14-like protein 3 (45 kDa secretory protein) (rsec45).
GN	Name=Sec143;
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC	Muridae; Murinae; Rattus.
OX	NCB1_TaxID=10116;
EN	[1]
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 5-11; 22-43; 52-58;
RC	174-184; 217-225; 231-253; 258-286; 269-279 AND 383-396.
RC	TISSUE=Olfactory epithelium;

Search completed: May 2, 2006, 18:38:13
Job time : 235 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 18:38:32 ; Search time 47 Seconds
(without alignments)
664.924 Million cell updates/sec

Title: US-10-696-699A-2
Perfect score: 2018
Sequence: 1 MSRGVGLSPRQKALAKPE.....LLPKASERKMQLGAGTPK 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/H COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1272	63.0	391	2	US-10-104-047-2798
2	1118	55.4	360	2	US-10-104-047-3609
3	394	19.5	723	2	US-09-949-016-9810
4	344.5	17.1	308	2	US-09-270-767-33202
5	344.5	17.1	308	2	US-09-270-767-48419
6	197	9.8	304	2	US-09-538-092-613
7	181.5	9.0	293	2	US-09-248-796A-18524
8	181	9.0	278	2	US-08-557-614-2
9	181	9.0	278	2	US-08-557-614-23
10	144	7.1	330	2	US-09-270-767-43290
11	124	6.1	301	2	US-09-270-767-43653
12	115.5	5.7	341	2	US-09-244-805-61
13	114.5	5.7	136	2	US-09-270-767-32985
14	114.5	5.7	136	2	US-09-270-767-48202
15	113	5.6	342	2	US-09-244-805-62
16	111	5.5	316	2	US-09-270-767-43378
17	111	5.5	778	2	US-09-270-767-45008
18	110	5.5	317	2	US-09-949-016-5927
19	110	5.5	344	2	US-09-949-016-9647
20	110	5.5	433	2	US-09-507-765-30
21	110	5.5	433	2	US-09-507-765-31
22	109	5.4	274	2	US-09-270-767-32136
23	109	5.4	274	2	US-09-270-767-47353
24	109	5.4	823	2	US-09-270-767-42450
25	107	5.3	135	2	US-09-270-767-58728
26	107	5.3	135	2	US-09-270-767-60491
27	107	5.3	342	2	US-09-248-796A-15115

28	105.5	5.2	108	2	US-09-270-767-59035	Sequence 59035, A
29	102	5.1	433	2	US-09-248-796A-18046	Sequence 18046, A
30	100.5	5.0	319	2	US-09-270-767-32632	Sequence 32632, A
31	100.5	5.0	319	2	US-09-270-767-47849	Sequence 47849, A
32	98.5	4.9	736	1	US-08-788-892-2	Sequence 2, Appli
33	98.5	4.9	736	1	US-09-340-475-2	Sequence 2, Appli
34	97	4.8	341	1	US-08-530-566-7	Sequence 7, Appli
35	97	4.8	341	2	US-09-195-726-7	Sequence 7, Appli
36	97	4.8	341	2	US-09-067-755-7	Sequence 9, Appli
37	95	4.7	341	1	US-08-530-566-9	Sequence 9, Appli
38	95	4.7	341	2	US-09-195-726-9	Sequence 9, Appli
39	95	4.7	341	2	US-09-067-755-9	Sequence 3753, Ap
40	95	4.7	418	2	US-09-134-001C-3753	Sequence 37, Appl
41	95	4.7	1440	2	US-09-357-251-37	Sequence 3655, Ap
42	94	4.7	736	2	US-09-134-001C-3655	Sequence 10, Appl
43	94	4.7	762	1	US-08-907-166-10	Sequence 10, Appl
44	94	4.7	762	2	US-09-391-340-10	Sequence 4, Appli
45	93	4.6	593	1	US-08-900-927-4	

ALIGNMENTS

RESULT 1
US-10-104-047-2798
; Sequence 2798, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2798
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-104-047-2798

Query Match	63.0%;	Score 1272;	DB 2;	Length 391;
Best Local Similarity	62.0%;	Pred. No. 3.1e-138;		
Matches	232;	Conservative 58;	Mismatches 68;	Indels 16; Gaps 2;
Qy	20	EASTCRSRPCSGS-----	MWSSB--SKRTLTTS	SLAWOPPEVIOQLSGGM 63
Db	14	ETLTCRNPKTCSRGKSPQPTSCPHALRAWHMEFRKQDDL	DNIVTWQPPEVIOQLYDSGGL 73	
Qy	64	CGYDLGCPWYDIIGPKDAKGLLFSKODLARTKWRCELLLOE	CAHOTTKLGRKVVET 123	
Db	74	CGYDEGCPYVFNIIIGSLDPKGLLSASKQDMIRKRIKVC	ELLHHECELOTQKLGRIEM 133	
Qy	124	ITTIYCEGGLKHLWKPAVEAYGEFLCMFEENYPETLKR	LFVVKAPKLPFVAVNLIKPF 183	
Db	134	ALIVFDMEGLSKHLWKPAVEVYQQFPSILEANYPETL	KNLIVIRAPKLPFVAFNLVKSF 193	
Qy	184	LSBDRTRKKIMVLGANYKEVILLKHISPDQVPEVGGT	WTDPDGPNPKCSKKNYGGDI PRKY 243	
Db	194	MSEETRKRKIVILGDNWKQELTKFISPDQLPVEFGT	WTDPDGPNPKCLTKINYGGEVPSY 253	
Qy	244	YVDRQVQOQYEHSHVQISRGSSHOVEYELPFGCVLR	WQFMSDGDADVGGIFLTKTWGBERQ 303	
Db	254	YLCEQVRLQYEHTRSVGRSSLOVENEILPFGCVLR	WQFASDGGDIDGFGVFLKTMGEQQ 313	
Qy	304	RAGBMTVLNORNSHLVPEDGTLTCTSDPGIYVLR	FDNTYSFTTHAKKVNFTVLELLPDK 363	
Db	314	SAREMTVLFSQRNANMVPEDGSLTCLQAGVYVLR	FDNTYSRMVAKKLSYTVLELLPDK 373	
Qy	364	ASBEKMKQLGAGTP 377		

7.1%; Score 144; DB 2; Length 330;
Best Local Similarity 26.4%; Pred. No. 1.3e-07;

Matches 48; Conservative 27; Mismatches 53; Indels 54; Gaps 7;
QY 86 LLSASKQDLLRTKME-----CELLQECAH-----112
Db 121 LKYGAAACENIIPSKLRNVFEANTILNLPQDQGRRLVLEAGKWKQSPQVPLVDFRGI 180
QY 113 QTTKLGKRVETIT-----IYDCGLGLKHL--WKPAVEA-----YGFPLCMFEENYPE 159
Db 181 QLTVLGSMVEPPSQIGSVVILDMGLPLSHITQTFPSFAAMLLDYIQECIM-----233
QY 160 TLKRLFFVKAPKLPFVAYNLKPLSEDTTRKKIMVLGANYKEVLKXHSIPDQVPVBYGGT 219
Db 234 RLKAVHVNNSYIFNMLFAVKEFFIREKLKRIFFGKDYKS-LISHIAKALPPKYGGS 292
QY 220 MT 221
Db 293 AT 294

RESULT 11
US-09-270-767-43653
; Sequence 43653, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43653
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43653

Query Match 6.1%; Score 124; DB 2; Length 301;
Best Local Similarity 22.2%; Pred. No. 2.4e-05;
Matches 41; Conservative 31; Mismatches 71; Indels 42; Gaps 6;
QY 84 KGLLFSASKQDLLR-----TKMRECE-----LLQEC-----AHQT 114
Db 103 RGLLVEQKEKFKVGSVINVLKNCQKGRVLIWNCGLWDPDITSDMFRLMYVHLA 162
QY 115 TKLGR--KVETITIIYDCGLGLKHLWKPAVEAYGEFLCMFEENYPETLKLFLVVKAPKL 172
Db 163 AQLSEETQVRGVVCIIMDFEGLSMKQKALSPSPSKRLLTFTQEAMPLRMKEVHFVKQFFI 222
QY 173 FPVAYNLKPLSEDTTRKKIMVLGANYKEVLKXHSIPDQVPVBYGGTMDPDGNPKCKSK 232
Db 223 FNMVWSLFPKPVKQLNRMHFHGDMSKS-LQKFLDPSVLPAHYKGTL-----PA 271
QY 233 INYGG 237
Db 272 IDYGG 276

RESULT 12
US-09-244-805-61
; Sequence 61, Application US/09244805
; Patent No. 6699660
; GENERAL INFORMATION:
; APPLICANT: Worley, Paul F.
; APPLICANT: Lanahan, Anthony
; APPLICANT: Goetz, Bernard
; APPLICANT: Heimisch, Holger
; APPLICANT: Kuner, Rohini
; APPLICANT: Scheek, Sigrid
; APPLICANT: Nikolich, Karoly
; APPLICANT: Zhukovski, Eugene
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
; THEREFOR

FILE REFERENCE: 10496/004001
; CURRENT APPLICATION NUMBER: US/09/244,805
; CURRENT FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/074,518
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: 60/074,135
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)---(341)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-244-805-61

Query Match 5.7%; Score 115.5; DB 2; Length 341;
Best Local Similarity 22.5%; Pred. No. 0.00029;
Matches 50; Conservative 36; Mismatches 85; Indels 51; Gaps 9;
QY 43 LTTSLAWQPPEVIQVYLSGCMGYD-----LDGC-----PWYDIIGPKDAKGLLPSA 90
Db 73 LSTSL---DDAFLRLFRARKFDYDRALQLLVNHYGCRSRWPEVFNLRPSALKDVLNSG 129
QY 91 SKQDLLRTKMECELL-----LQSCAHTTKLGRKVT-----ITIIYD 129
Db 130 FLTVLPHTPRGCHVLCIRPDWIPSNYPITENIRAIYLTLEKLIQSEETQVNGVILAD 189
QY 130 CEGGLGLKHLWKPAVEAYGEFLC-----MFEENYPETLKLFLVVKAPKLFPVAYNLKPL 184
Db 190 YKGVSL-----KASHFGPFIAKVGILQDGPPIKAVHIVNEPRFKGIFAIKPL 244
QY 185 SEDTRKKIMVLGANYKEVLKXHSIPDQ--VPVEYGGTMDPD 224
Db 245 KEKIANRFFLHGSDDL---HTSLPRNLPKVEYGGTAGELD 283

RESULT 13
US-09-270-767-32985
; Sequence 32985, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32985
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32985

Query Match 5.7%; Score 114.5; DB 2; Length 136;
Best Local Similarity 24.6%; Pred. No. 8e-05;
Matches 33; Conservative 35; Mismatches 49; Indels 17; Gaps 6;
QY 240 PRKYVYRDQVQKQVEH---SVQISRGSSHOVEYEI-LFPGCVLRQWQMSDQADVGFGIF- 294
Db 1 PEELIYI-DQSSQSDRDFVEAQVPGDKLKLHFKVWVEEQKILSWERTFDYDIKFGIYS 59
QY 295 LKTKWGERQAGEMTEVLPNQYNSHLVPEBDGTLTCDSPDGGIYVLRFPNTYVSFIHAKKNF 354
Db 60 VDDKTGKRSEVPVLTGYVNE-----MDEIGVISTRPNTTYTVVDFNSASAYLSRKKLY 113
QY 355 TVEVLLPDKASEEK 368
Db 114 WVDLI-----SEEE 122

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 18:49:42 ; Search time 165 Seconds
(without alignments)
957.209 Million cell updates/sec

Title: US-10-696-699A-2
Perfect score: 2018
Sequence: 1 MSRGVDSLSPRQKEALAKPE.....LLPDKASEBKKQLGAGTPK 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2018	100.0	378	4	US-10-419-629-2
2	2018	100.0	378	4	US-10-696-699A-2
3	1827.5	90.6	403	4	US-10-220-475A-4
4	1827.5	90.6	403	4	US-10-419-629-4
5	1827.5	90.6	403	4	US-10-072-012-448
6	1827.5	90.6	403	4	US-10-696-699A-4
7	1827.5	90.6	403	5	US-10-723-860-1875
8	1827.5	90.6	403	5	US-10-732-923-22953
9	1827.5	90.6	403	5	US-10-732-923-22954
10	1827.5	90.6	403	5	US-10-631-467-841
11	1827.5	90.6	403	5	US-10-745-237-294
12	1724.5	85.5	403	4	US-10-072-012-450
13	1724.5	85.5	403	5	US-10-732-923-22775
14	1722.5	85.4	403	4	US-10-220-475A-2
15	1722.5	85.4	403	4	US-10-072-012-449
16	1722.5	85.4	403	5	US-10-732-923-22778
17	1685.5	83.5	403	5	US-10-732-923-22958
18	1616.5	80.1	387	5	US-10-732-923-22957
19	1548	76.7	293	4	US-10-696-699A-17
20	1428	70.8	400	4	US-10-220-475A-1
21	1428	70.8	400	5	US-10-732-923-22956
22	1423.5	70.5	419	5	US-10-450-763-43464
23	1413	70.0	400	4	US-10-072-012-451
24	1413	70.0	400	5	US-10-732-923-22780
25	1309.5	64.9	403	5	US-10-732-923-22774
26	1304.5	64.6	406	4	US-10-220-475A-18
27	1304.5	64.6	406	4	US-10-072-012-447

28	1304.5	64.6	406	5	US-10-732-923-22950	Sequence 22950, A
29	1272	63.0	391	4	US-10-104-047-2798	Sequence 2798, Ap
30	1254	62.1	416	4	US-10-072-012-124	Sequence 124, App
31	1227.5	60.8	415	4	US-10-072-012-122	Sequence 122, App
32	1200.5	59.5	502	5	US-10-450-763-43046	Sequence 43046, A
33	1194	59.2	226	4	US-10-696-699A-17	Sequence 17, Appl
34	1148.5	56.9	405	4	US-10-072-012-126	Sequence 126, App
35	1118	55.4	360	4	US-10-104-047-3609	Sequence 3609, Ap
36	809	40.1	300	4	US-10-336-472-46	Sequence 46, Appl
37	804	39.8	151	4	US-10-696-699A-19	Sequence 43044, A
38	752	37.3	240	5	US-10-450-763-43044	Sequence 48, Appl
39	689	34.1	268	4	US-10-336-472-48	Sequence 22946, A
40	661.5	32.8	343	5	US-10-732-923-22946	Sequence 22959, A
41	518	25.7	385	5	US-10-732-923-22959	Sequence 22977, A
42	439	21.8	407	5	US-10-732-923-22977	Sequence 100, App
43	439	21.8	407	5	US-10-745-237-100	Sequence 33039, A
44	435	21.6	385	6	US-11-097-143-33039	Sequence 6, Appli
45	433	21.5	696	4	US-10-478-245-6	

ALIGNMENTS

RESULT 1

US-10-419-629-2
; Sequence 2, Application US/10419629
; Publication No. US20040023915A1
; GENERAL INFORMATION:
; APPLICANT: Sanders, Bob G.
; APPLICANT: Kline, Kimberly
; APPLICANT: Yu, Weiping
; APPLICANT: Liu, Hui
; APPLICANT: Hantash, Feras
; TITLE OF INVENTION: Tocopherol Associated Protein and Uses Thereof
; FILE REFERENCE: D6453
; CURRENT APPLICATION NUMBER: US/10/419,629
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/373,870
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: P2P polypeptide
US-10-419-629-2

Query Match 100.0%; Score 2018; DB 4; Length 378;
Best Local Similarity 100.0%; Pred. No. 8e-194;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSRGVDSLSPRQKEALAKPEASTCRSRPCSGMWSSESKRTLTTSLAWQPVEVQYLS	60
Db	1	MSRGVDSLSPRQKEALAKPEASTCRSRPCSGMWSSESKRTLTTSLAWQPVEVQYLS	60
QY	61	GMCGVDLDCGPVWYDIIGPKDAKGLLFSASKODLLATKRECELLILOCAHQTTKLGK	120
Db	61	GMCGVDLDCGPVWYDIIGPKDAKGLLFSASKODLLATKRECELLILOCAHQTTKLGK	120
QY	121	VEITITTYDCGGLGKHLWKPAVEAYGEFLCMPEENYPETLKRLLFVVKAPKLPVAVNLI	180
Db	121	VEITITTYDCGGLGKHLWKPAVEAYGEFLCMPEENYPETLKRLLFVVKAPKLPVAVNLI	180
QY	181	KPFLSETRKKIIVLGANYKEVLLKHISPDQVPVEYGGTMTDPDGNPKCKSKINYGDI	240
Db	181	KPFLSETRKKIIVLGANYKEVLLKHISPDQVPVEYGGTMTDPDGNPKCKSKINYGDI	240
QY	241	RYKYVRDQVKQVEHSHVQISRGSSHQVEYIELFPGCVLRWQFMSDGDVGFGLKTKMG	300
Db	241	RYKYVRDQVKQVEHSHVQISRGSSHQVEYIELFPGCVLRWQFMSDGDVGFGLKTKMG	300

QY 301 ERQRAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGIVVLPFDNTYSFIHAKKNFTVEVLL 360
|||||
Db 301 ERQRAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGIVVLPFDNTYSFIHAKKNFTVEVLL 360
QY 361 PDKASEKMKQLGAGTPK 378
|||||
Db 361 PDKASEKMKQLGAGTPK 378

RESULT 2

US-10-696-699A-2
; Sequence 2, Application US/10696699A
; Publication No. US20040152883A1
; GENERAL INFORMATION:
; APPLICANT: Sanders, Bob G.
; APPLICANT: Kline, Kimberly
; APPLICANT: Yu, Weiping
; APPLICANT: Liu, Hui
; APPLICANT: Hantash, Feras
; TITLE OF INVENTION: Tocopherol Associated Protein and Uses Thereof
; FILE REFERENCE: D6453CIP
; CURRENT APPLICATION NUMBER: US/10/696,699A
; PRIOR FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 10/419,629
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: TAP-38 polypeptide
US-10-696-699A-2

Query Match 100.0%; Score 2018; DB 4; Length 378;
Best Local Similarity 100.0%; Pred. No. 8.8e-194;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGRVGDLSPROKEALAKPEASTCRSRRCSGSMSSSESKRTLTTSIAWQPPEVIQQYLS 60
|||||
Db 1 MSGRVGDLSPROKEALAKPEASTCRSRRCSGSMSSSESKRTLTTSIAWQPPEVIQQYLS 60
QY 61 GWCYVDLDCGPWYDIIGPKDAKGLLFSASQDILLRTKMRCELLLOQCAHQTTLGRK 120
|||||
Db 61 GWCYVDLDCGPWYDIIGPKDAKGLLFSASQDILLRTKMRCELLLOQCAHQTTLGRK 120
QY 121 VETITIIYDCEGLGLKHLKHPAVEAYGEFLCMFEENYPTLKRLLFVVKAPKLPVAYNLI 180
|||||
Db 121 VETITIIYDCEGLGLKHLKHPAVEAYGEFLCMFEENYPTLKRLLFVVKAPKLPVAYNLI 180
QY 181 KPFLSDTRKKIMVLGANYKEVLLKHISPDQVPVEYGGTMDPDGNPKCKSKINYGDDIP 240
|||||
Db 181 KPFLSDTRKKIMVLGANYKEVLLKHISPDQVPVEYGGTMDPDGNPKCKSKINYGDDIP 240
QY 241 RXYVVDQVKQYEHVSQVSRGSSHQVEYELLFPGCVLRWQFMSDQADVGFGLTKRMG 300
|||||
Db 241 RXYVVDQVKQYEHVSQVSRGSSHQVEYELLFPGCVLRWQFMSDQADVGFGLTKRMG 300
QY 301 ERQRAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGIVVLPFDNTYSFIHAKKNFTVEVLL 360
|||||
Db 301 ERQRAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGIVVLPFDNTYSFIHAKKNFTVEVLL 360
QY 361 PDKASEKMKQLGAGTPK 378
|||||
Db 361 PDKASEKMKQLGAGTPK 378

RESULT 3

US-10-220-475A-4
; Sequence 4, Application US/10220475A
; Publication No. US20040023227A1
; GENERAL INFORMATION:

; APPLICANT: INOUE, KEIZO
; APPLICANT: ARAI, HIROYUKI
; APPLICANT: ARITA, MAKOTO
; TITLE OF INVENTION: DNA ENCODING SQUALENE EPOXIDASE-PROMOTING FACTOR
; FILE REFERENCE: 04853.0096-00000
; CURRENT APPLICATION NUMBER: US/10/220,475A
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: PCT/JF01/01592
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: JP 2000-57743
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-475A-4

Query Match 90.6%; Score 1827.5; DB 4; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 19; Mismatches 18; Indels 25; Gaps 2;
QY 1 MSGRVGDLSPROKEALAK+++++++PEASTCRSRRCSGSMSSSES+++++ 39
|||||
Db 1 MSGRVGDLSPROKEALAKFRENVDVLPALPNPDDYFLRLWLARSFDLQKSEAMLRKHV 60
QY 40 ++++++SLAWQPPEVIQQYLSGGMCGYDLDCGVPWYDIIGPKDAKGLLFSASKQDL 95
|||||
Db 61 EFRKQKIDNLTISWQPPEVIQQYLSGGMCGYDLDCGVPWYDIIGPKDAKGLLFSASKQDL 120
QY 96 LRTKMRCELLLOQCAHQTTLGRKVTITIIYDCEGLGLKHLKHPAVEAYGEFLCMFEE 155
|||||
Db 121 LRTKMRCELLLOQCAHQTTLGRKVTITIIYDCEGLGLKHLKHPAVEAYGEFLCMFEE 180
QY 156 NYPETLKRLLFVVKAPKLPVAYNLIKPFLSDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
|||||
Db 181 NYPETLKRLLFVVKAPKLPVAYNLIKPFLSDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
QY 216 YGGTMDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQVSRGSSHQVEYELLFPG 275
|||||
Db 241 YGGTMDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQVSRGSSHQVEYELLFPG 300
QY 276 CVLRWQFMSDQADVGFGLTKRMGERORAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGI 335
|||||
Db 301 CVLRWQFMSDQADVGFGLTKRMGERORAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGI 360
QY 336 VYLRFDNTYSFIHAKKNFTVEVLLPDKASEKMKQLGAGTPK 378
|||||
Db 361 VYLRFDNTYSFIHAKKNFTVEVLLPDKASEKMKQLGAGTPK 403

RESULT 4

US-10-419-629-4
; Sequence 4, Application US/10419629
; Publication No. US20040023915A1
; GENERAL INFORMATION:
; APPLICANT: Sanders, Bob G.
; APPLICANT: Kline, Kimberly
; APPLICANT: Yu, Weiping
; APPLICANT: Liu, Hui
; APPLICANT: Hantash, Feras
; TITLE OF INVENTION: Tocopherol Associated Protein and Uses Thereof
; FILE REFERENCE: D6453
; CURRENT APPLICATION NUMBER: US/10/419,629
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US 60/373,870
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 4
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: TAP-46 polypeptide
US-10-419-629-4

Query Match          90.6%; Score 1827.5; DB 4; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;

Qy 1 MSGRVGDLSPRQKALAK-----PEASTCSRPRRCSSGMWSES----- 39
Db 1 MSGRVGDLSPRQKALAKRFENVQDVLPALENPDYFLRLWLRLARSFDLQKSEAMLRKHV 60
Qy 40 ----KRTLTTSLAWQPPREVIQQYLSGGCGYDLGCPWYDIIGPKDAKGLLFSASKQDL 95
Db 61 EFRQKQIDNIIISWQPPREVIQQYLSGGCGYDLGCPWYDIIGPLDAKGLLFSASKQDL 120
Qy 96 LRTKMRCELLLQECAHQTTKLGRLKRVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMPEE 155
Db 121 LRTKMRCELLLQECAHQTTKLGRLKRVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMPEE 180
Qy 156 NYPETLKLRLFVVKAPKLPFVAYNLIKPLFSDTRKKIMVLGANYKEVILLKHISPDQVPE 215
Db 181 NYPETLKLRLFVVKAPKLPFVAYNLIKPLFSDTRKKIMVLGANYKEVILLKHISPDQVPE 240
Qy 216 YGGTWTDPDGNPKCKSKINYGGDIPRKYYVRDQVKQYEHVSQVISRGSSHQVEYELLFPG 275
Db 241 YGGTWTDPDGNPKCKSKINYGGDIPRKYYVRDQVKQYEHVSQVISRGSSHQVEYELLFPG 300
Qy 276 CVLRWQFMSDGNADVGFGLFKLTQMGERQAGEMTEVLPNQRVNSHLVPEDGTLTCSDFGI 335
Db 301 CVLRWQFMSDGNADVGFGLFKLTQMGERQAGEMTEVLPNQRVNSHLVPEDGTLTCSDFGI 360
Qy 336 YVLRFDNTYSFTHAKKNFTVEVLLPDKASBEKKQLGAGTPK 378
Db 361 YVLRFDNTYSFTHAKKNFTVEVLLPDKASBEKKQLGAGTPK 403

```

RESULT 5
US-10-072-012-448
; Sequence 448, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Pattursajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514

```
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 448
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-448

Query Match          90.6%; Score 1827.5; DB 4; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2

Qy      1 MSGRVGDLSPRQKEALAK-----PEASTCRSRRCSSGSMWSSES----- 39
        |||
Db      1 MSGRVGDLSPRQKEALAKFRENVQDVLPAFPNDPDYFLLRWLRLARSFDLQKSEAMLRKHV 60
        |||

Qy      40 ----KRTLITSLAWQPPEVIQQYLXSGMGCGYDLGGCPWVYDIIGPKDAKGLLFSASKODL 95
        :::::
Db      61 EFRKQKDIDNIISWQPPEVIQQYLXSGMGCGYDLGGCPWVYDIIGPLDKAGLLFSASKODL 120

Qy      96 LRTKWRECELLQBCAHOHTTKLGKRKVETITIIYDC EGLGLKHLWKPAVEAYGEFLCMFEE 155
        :
Db      121 LRTKWRECELLQBCAHOHTTKLGKRKVETITIIYDC EGLGLKHLWKPAVEAYGEFLCMFEE 180

Qy      156 NYPETLKRLFVVKAPKLPFVAYNLIKPLSDSTRKKIMVLGCANYKEVULLKHISPDPQVPE 215
        :
Db      181 NYPETLKRLFVVKAPKLPFVAYNLIKPLSDSTRKKIMVLGCANYKEVULLKHISPDPQVPE 240

Qy      216 YGGTMTDPDGNGPKCKSKINYGDDIPRKYYVRDVKQOQYEHVSVOISRGSSSHQVEYIILPFG 275
        :
Db      241 YGGTMTDPDGNGPKCKSKINYGDDIPRKYYVRDVKQOQYEHVSVOISRGSSSHQVEYIILPFG 300

Qy      276 CVLRWFQMSDGDADVGFGIFLTKNQRORAGEMTEVLPNQRYNSHLPVEDGTTLTCSDPGI 335
        :
Db      301 CVLRWFQMSDGDADVGFGIFLTKNQRORAGEMTEVLPNQRYNSHLPVEDGTTLTCSDPGI 360

Qy      336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEBKMQLGAGTPK 378
        |||
Db      361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEBKMQLGAGTPK 403


RESULT 6
US-10-696-699A-4
; Sequence 4, Application US/10696699A
; Publication No. US20040152883A1
; GENERAL INFORMATION:
; APPLICANT: Sanders, Bob G.
; APPLICANT: Kline, Kimberly
; APPLICANT: Yu, Weiping
; APPLICANT: Liu, Hui
; APPLICANT: Hantash, Reras
; TITLE OF INVENTION: Tocopherol Associated Protein and Uses Thereof
; FILE REFERENCE: D6453CIP
; CURRENT APPLICATION NUMBER: US/10/696,699A
; CURRENT FILING DATE: 2003-10-29
```

```

; PRIOR APPLICATION NUMBER: US 10/419,629
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 4
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: TAP-46 polypeptide
US-10-696-699a-4

Query Match          90.6%; Score 1827.5; DB 4; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPRQKEALAKFRENVDVLPALPNPDDYFLLRWLRARSFDLQKSEAMLRKHV 60
Db 1 MSGRVGDLSPRQKEALAKFRENVDVLPALPNPDDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTTSIAWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
Db 40 ----KRTLTTSIAWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
QY 61 EFRKQKIDNIISWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120
Db 61 EFRKQKIDNIISWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120
QY 96 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
Db 96 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
QY 121 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
Db 121 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
QY 156 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
Db 156 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
QY 181 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
Db 181 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
QY 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 275
Db 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 275
QY 241 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 300
Db 241 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 300
QY 276 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 335
Db 276 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 335
QY 301 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 360
Db 301 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 360
QY 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 378
Db 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 378
QY 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 403
Db 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 403

RESULT 7
US-10-723-860-1875
; Sequence 1875, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1875
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1875

Query Match          90.6%; Score 1827.5; DB 5; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPRQKEALAK-----PEASTCRSRRPCSGSMWSSES----- 39
Db 1 MSGRVGDLSPRQKEALAKFRENVDVLPALPNPDDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTTSIAWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
Db 40 ----KRTLTTSIAWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
QY 61 EFRKQKIDNIISWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120
Db 61 EFRKQKIDNIISWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120
QY 96 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
Db 96 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
QY 121 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
Db 121 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
QY 156 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
Db 156 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
QY 181 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
Db 181 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
QY 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 275
Db 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 275
QY 241 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 300
Db 241 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 300
QY 276 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 335
Db 276 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 335
QY 301 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 360
Db 301 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 360
QY 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 378
Db 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 378
QY 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 403
Db 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKMQLGAGTPK 403

RESULT 8
US-10-732-923-22953
; Sequence 22953, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 22953
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-22953

Query Match          90.6%; Score 1827.5; DB 5; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPRQKEALAK-----PEASTCRSRRPCSGSMWSSES----- 39
Db 1 MSGRVGDLSPRQKEALAKFRENVDVLPALPNPDDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTTSIAWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
Db 40 ----KRTLTTSIAWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPKDAKGLLFSASKQDL 95
QY 61 EFRKQKIDNIISWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120
Db 61 EFRKQKIDNIISWQPPEVIOQYLSGGMCGYDLDCGCPVWYDIIGPLDAKGLLFSASKQDL 120
QY 96 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
Db 96 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 155
QY 121 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
Db 121 LRTKMRECELLLOECAHQTTLGRKVETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE 180
QY 156 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
Db 156 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
QY 181 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
Db 181 NYPETLKLFLVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
QY 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 275
Db 216 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 275
QY 241 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 300
Db 241 YGGTMTDPDGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELPPG 300
QY 276 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 335
Db 276 CVLRQWFMSDGDADVGFGLFKTKMGERORAGEMTEVLPNQRYNHSLVPEDGTLTCSDFGI 335
```


Db 301 CVLRWQFMSDGDVGFGLFKTKMGERQORAGEMTEVLPNQRYNSHLVDPDGTLTCSDPGI 360
QY 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKKQKGAGTPK 378
Db 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKKQKGAGTPK 403

RESULT 9

US-10-732-923-22954
; Sequence 22954, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 22954
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-732-923-22954

Query Match 90.6%; Score 1827.5; DB 5; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPROKEALAK-----PEASTCRSRPPCGSGMWSSES----- 39
Db 1 MSGRVGDLSPROKEALAKFRENVDVLPALPNPDDYFLLRLWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTSLAWQPPEVIQQYLSGCMCGYLDGCPVWYDIIGPKDAKGLLFSASKQDL 95
Db 61 EPRKQKIDNIISWQPPEVIQQYLSGCMCGYLDGCPVWYDIIGPKDAKGLLFSASKQDL 120
QY 96 LRTKMRCELLQECACHQTTKLGKRVETITIIYDCBGLGKHLWKPAVEAYGFLCMFEE 155
Db 121 LRTKMRCELLQECACHQTTKLGKRVETITIIYDCBGLGKHLWKPAVEAYGFLCMFEE 180
QY 156 NYPETLKRFLVVKAPKLPVAVYNLIKPFLESDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
Db 181 NYPETLKRFLVVKAPKLPVAVYNLIKPFLESDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
QY 216 YGGTMTDPGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELLFPG 275
Db 241 YGGTMTDPGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELLFPG 300
QY 276 CVLRWQFMSDGDVGFGLFKTKMGERQORAGEMTEVLPNQRYNSHLVDPDGTLTCSDPGI 335
Db 301 CVLRWQFMSDGDVGFGLFKTKMGERQORAGEMTEVLPNQRYNSHLVDPDGTLTCSDPGI 360
QY 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKKQKGAGTPK 378
Db 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKKQKGAGTPK 403

RESULT 10

US-10-631-467-841
; Sequence 841, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for broncheal asthma, or chronic obstructive P
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312

; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 841
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-841

Query Match 90.6%; Score 1827.5; DB 5; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPROKEALAK-----PEASTCRSRPPCGSGMWSSES----- 39
Db 1 MSGRVGDLSPROKEALAKFRENVDVLPALPNPDDYFLLRLWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTSLAWQPPEVIQQYLSGCMCGYLDGCPVWYDIIGPKDAKGLLFSASKQDL 95
Db 61 EPRKQKIDNIISWQPPEVIQQYLSGCMCGYLDGCPVWYDIIGPKDAKGLLFSASKQDL 120
QY 96 LRTKMRCELLQECACHQTTKLGKRVETITIIYDCBGLGKHLWKPAVEAYGFLCMFEE 155
Db 121 LRTKMRCELLQECACHQTTKLGKRVETITIIYDCBGLGKHLWKPAVEAYGFLCMFEE 180
QY 156 NYPETLKRFLVVKAPKLPVAVYNLIKPFLESDTRKKIMVLGANYKEVLLKHISPDQVPVE 215
Db 181 NYPETLKRFLVVKAPKLPVAVYNLIKPFLESDTRKKIMVLGANYKEVLLKHISPDQVPVE 240
QY 216 YGGTMTDPGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELLFPG 275
Db 241 YGGTMTDPGNPKCKSKINYGDDIPRKYVVRDQVKQYEHVSQISRGSSHQVEYELLFPG 300
QY 276 CVLRWQFMSDGDVGFGLFKTKMGERQORAGEMTEVLPNQRYNSHLVDPDGTLTCSDPGI 335
Db 301 CVLRWQFMSDGDVGFGLFKTKMGERQORAGEMTEVLPNQRYNSHLVDPDGTLTCSDPGI 360
QY 336 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKKQKGAGTPK 378
Db 361 YVLRFDNTYSFIHAKKNFTVEVLLPDKASEEKKQKGAGTPK 403

RESULT 11

US-10-745-237-294
; Sequence 294, Application US/10745237
; Publication No. US20050227301A1
; GENERAL INFORMATION:
; APPLICANT: Cyclacel Limited
; APPLICANT: Glover, David
; APPLICANT: Bell, Graham
; APPLICANT: Frenz, Lisa
; APPLICANT: Midgley, Carol
; TITLE OF INVENTION: Cell Cycle Progression Proteins
; FILE REFERENCE: P0158190 CYK
; CURRENT APPLICATION NUMBER: US/10/745,237
; CURRENT FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: US 60/439,123
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: US 60/468,402
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 294
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: O76054
US-10-745-237-294

Query Match 90.6%; Score 1827.5; DB 5; Length 403;
Best Local Similarity 87.1%; Pred. No. 1.4e-174;
Matches 351; Conservative 9; Mismatches 18; Indels 25; Gaps 2;


```
QY 1 MSRGVGDLSRQKEALAK-----PEASTCSRPPCSGSMWSSES----- 39
DB 1 MSRGVGDLSRQKEALAKFRENVDVLPPLNPDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTTSLAWQPPVPIQQYLSGCMGCGYDLGCPVWYDIIGPKDAKGLLFSASKQDL 95
DB 61 EFRKQKDIKISWQPPVPIQQYLSGRCGYDLGCPVWYDIIGPLDAKGLLFSASKQDL 120
QY 96 LRTKMRCELLQCEAHQTTKLGKRVETITIIYDCBGLGLKHLWKPAVAYGEFLTMFEE 155
DB 121 LRTKMRCELLQCEIQTQTKLGKIKETITIIYDCBGLGLKHLWKPAVAYGEFLTMFEE 180
QY 156 NYPETLKRFLFVKAPKLPFVAYNLIKPFLSEDRKKIMVLGANYKEVLLKHSPPDQVPE 215
DB 181 NYPETLKRFLFVKAPKLPFVAYNLIKPFLSEDRKKIMVLGANYKEVLLKHSPPDQVPE 240
QY 216 YGGTMTDPDGNPKCKSKINYGDIIPKYYVVRDQVQKQYEHVSQISRGSSHQVEYELFPFG 275
DB 241 YGGTMTDPDGNPKCKSKINYGDIIPKYYVVRDQVQKQYEHVSQISRGSSHQVEYELFPFG 300
QY 276 CVLRWQFMSDGVGFGIFLTKMGERQORAGEMTEVLPNQRYNHSHLVPEDEGTLTCSDPGI 335
DB 301 CVLRWQFMSGSDVGFIFLTKMGERQORAGEMTEVLPNQRYNHSHLVPEDEGTLTCSDEPGI 360
QY 336 YVLRFDNTYSFIHAKKNVFTVEVLLPDKASEEKKQLGAGTPK 378
DB 361 YVLRFDNTYSFIHAKKNVFTVEVLLPDKAAEEKLNQOQNAVTEK 403
```

RESULT 14

```
US-10-220-475A-2
; Sequence 2, Application US/10220475A
; Publication No. US2004003227A1
; GENERAL INFORMATION:
; APPLICANT: INOUE, KEIZO
; APPLICANT: ARAI, HIROYUKI
; APPLICANT: ARITA, MAKOTO
; TITLE OF INVENTION: DNA ENCODING SQUALENE EPOXIDASE-PROMOTING FACTOR
; FILE REFERENCE: 04853.0096-00000
; CURRENT APPLICATION NUMBER: US/10/220,475A
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: PCT/JP01/01592
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: JP 2000-57743
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-220-475A-2
```

```
Query Match 85.4%; Score 1722.5; DB 4; Length 403;
Best Local Similarity 81.1%; Pred. No. 5.2e-164;
Matches 327; Conservative 25; Mismatches 26; Indels 25; Gaps 2;
```

```
QY 1 MSRGVGDLSRQKEALAK-----PEASTCSRPPCSGSMWSSES----- 39
DB 1 MSRGVGDLSRQKEALAKFRENVDVLPALPNPDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ----KRTLTTSLAWQPPVPIQQYLSGCMGCGYDLGCPVWYDIIGPKDAKGLLFSASKQDL 95
DB 61 EFRKQKDIKISWQPPVPIQQYLSGRCGYDLGCPVWYDIIGPLDAKGLLFSASKQDL 120
QY 96 LRTKMRCELLQCEAHQTTKLGKRVETITIIYDCBGLGLKHLWKPAVAYGEFLTMFEE 155
DB 121 LRTKMRCELLQCEIQTQTKLGKIKETITIIYDCBGLGLKHLWKPAVAYGEFLTMFEE 180
QY 156 NYPETLKRFLFVKAPKLPFVAYNLIKPFLSEDRKKIMVLGANYKEVLLKHSPPDQVPE 215
DB 181 NYPETLKRFLFVKAPKLPFVAYNLIKPFLSEDRKKIMVLGANYKEVLLKHSPPDQVPE 240
```

```
QY 216 YGGTMTDPDGNPKCKSKINYGDIIPKYYVVRDQVQKQYEHVSQISRGSSHQVEYELFPFG 275
DB 241 YGGTMTDPDGNPKCKSKINYGDIIPKYYVVRDQVQKQYEHVSQISRGSSHQVEYELFPFG 300
QY 276 CVLRWQFMSDGVGFGIFLTKMGERQORAGEMTEVLPNQRYNHSHLVPEDEGTLTCSDPGI 335
DB 301 CVLRWQFMSGSDVGFIFLTKMGERQORAGEMTEVLPNQRYNHSHLVPEDEGTLTCSDEPGI 360
QY 336 YVLRFDNTYSFIHAKKNVFTVEVLLPDKASEEKKQLGAGTPK 378
DB 361 YVLRFDNTYSFIHAKKNVFTVEVLLPDKAAEEKLNQOQNAVTEK 403

RESULT 15
US-10-072-012-449
; Sequence 449, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Molenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsebrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 449
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Rattus norvegicus
```

US-10-072-012-449

```
Query Match      85.4%; Score 1722.5; DB 4; Length 403;
Best Local Similarity 81.1%; Pred. No. 5.2e-164;
Matches 327; Conservative 25; Mismatches 26; Indels 25; Gaps 2;

QY 1 MSGRVGDLSPQKEALAK-----PEASTCRSRRPCSGSMWSSES----- 39
Db 1 MSGRVGDLSPQKEALAKFENVQDVLPAIPNPDYFLLRWLRARSFDLQKSEAMLRKHV 60
QY 40 ---KRTLTTSIAWQPPEVIQQYLSGGMCGYDLGCPVWYDIIGPKDAKGLLFSASKODL 95
Db 61 EFRKQKDIKIISWQPPEVIQQYLSGGRCGYDLGCPVWYDIIGPLDAKGLLFSASKODL 120
QY 96 LRTKMRCELLQECACHQTTKLGRKVETITIIYDCEGLGLKHLWKPVEAYGEFLCMFEE 155
Db 121 LRTKMRDCELLQECTQQTKLGGKTIETITMIYDCEGLGLKHLWKPVEAYGEFLTMFEE 180
QY 156 NYPETLKRLFPVVKAPKLPVAYNLIKPFLSEDTTRKKIMVLGANYKEVLLKHISPDQVPVE 215
Db 181 NYPETLKRLFPVVKAPKLPVAYNLIKPFLSEDTTRKKIMVLGANYKEVLLKHISPDQLPVE 240
QY 216 YGGTMTDPDGNPKCKSKINYGDDIPKYYVRDQVQOYEHVSQISRGSSHQVEYELPPG 275
Db 241 YGGTMTDPDGNPKCKSKINYGDDIPKYYVRDQVQOYEHVSQISRGSSHQVEYELPPG 300
QY 276 CVLRWQFMSDGADVGFGI FLKTKMGERQORAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGI 335
Db 301 CVLRWQFMSGSDVGFGI FLKTKMGERQORAGEMTEVLPNQRYNSHMVPEDGTLTCEPGI 360
QY 336 YVLRFDNTYSFIHAKKVNTTVEVLLPDKASEKMKQLGAGTPK 378
Db 361 YVLRFDNTYSFIHAKKVSTVEVLLPDKAAEEKLNQOGAVTPK 403
```

Search completed: May 2, 2006, 18:53:26
Job time : 166 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 18:50:52 ; Search time 27 Seconds
(without alignments)
636.690 Million cell updates/sec

Title: US-10-696-699A-2

Perfect score: 2018

Sequence: 1 MSGRVGDLSPRQKEALAKE.....LLPKASBEKMKQLGAGTPK 378

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 4547862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SIDSS5/ptodata/2/pubpaa/US06_NEW_PUB.pep1.*
- 2: /SIDSS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /SIDSS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /SIDSS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /SIDSS5/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /SIDSS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 7: /SIDSS5/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
- 8: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 9: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
- 10: /SIDSS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 11: /SIDSS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
- 12: /SIDSS5/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1827.5	90.6	403	9	US-10-528-031-7
2	1272	63.0	391	11	US-11-072-512-2798
3	1118	55.4	360	11	US-11-072-512-3609
4	233	11.5	463	11	US-11-096-568A-7798
5	233	11.5	479	11	US-11-096-568A-7797
6	233	11.5	483	11	US-11-096-568A-7796
7	221.5	11.0	478	11	US-11-096-568A-28170
8	221.5	11.0	517	11	US-11-096-568A-28169
9	221.5	11.0	558	11	US-11-096-568A-28168
10	216	10.7	409	11	US-11-096-568A-32080
11	216	10.7	433	11	US-11-096-568A-32079
12	214	10.6	511	11	US-11-096-568A-27441
13	214	10.6	551	11	US-11-096-568A-27440
14	214	10.6	560	11	US-11-096-568A-27439
15	213	10.6	394	11	US-11-096-568A-9415
16	213	10.6	423	11	US-11-096-568A-9414
17	213	10.6	449	11	US-11-096-568A-9413
18	206	10.2	419	11	US-11-096-568A-26554
19	206	10.2	443	11	US-11-096-568A-26553
20	206	10.2	463	11	US-11-096-568A-18643
21	206	10.2	485	11	US-11-096-568A-26552

22	204	10.1	341	11	US-11-096-568A-32081	Sequence 32081, A
23	199.5	9.9	546	11	US-11-096-568A-27860	Sequence 27860, A
24	199.5	9.9	550	11	US-11-096-568A-27859	Sequence 27859, A
25	199.5	9.9	554	11	US-11-096-568A-27858	Sequence 27858, A
26	196	9.7	364	11	US-11-087-099-7476	Sequence 7476, Ap
27	193.5	9.6	313	11	US-11-096-568A-16949	Sequence 16949, A
28	193.5	9.6	520	11	US-11-096-568A-16948	Sequence 16948, A
29	186.5	9.2	302	11	US-11-087-099-6162	Sequence 6162, Ap
30	176.5	8.7	497	11	US-11-087-099-5884	Sequence 5884, Ap
31	176	8.7	137	11	US-11-096-568A-26460	Sequence 26460, A
32	176	8.7	156	11	US-11-096-568A-26459	Sequence 26459, A
33	173.5	8.6	430	11	US-11-096-568A-18644	Sequence 18644, A
34	158	7.8	247	11	US-11-087-099-5238	Sequence 5238, Ap
35	149	7.4	307	11	US-11-096-568A-18887	Sequence 18887, A
36	149	7.4	309	11	US-11-096-568A-18886	Sequence 18886, A
37	142.5	7.1	208	11	US-11-096-568A-13413	Sequence 13413, A
38	142.5	7.1	255	11	US-11-096-568A-13412	Sequence 13412, A
39	142.5	7.1	289	11	US-11-096-568A-13411	Sequence 13411, A
40	142	7.0	259	11	US-11-087-099-1922	Sequence 1922, Ap
41	141	7.0	254	11	US-11-087-099-10781	Sequence 10781, A
42	141	7.0	261	11	US-11-087-099-1647	Sequence 1647, Ap
43	140.5	7.0	255	11	US-11-087-099-308	Sequence 308, App
44	140	6.9	256	11	US-11-087-099-3617	Sequence 3617, Ap
45	140	6.9	256	11	US-11-087-099-10493	Sequence 10493, A

ALIGNMENTS

RESULT 1

US-10-528-031-7
; Sequence 7, Application US/10528031
; Publication No. US20050262577A1
; GENERAL INFORMATION:
; APPLICANT: ORIDIS BIOMED Forschungs- und Entwicklungs GmbH
; APPLICANT: Guelly, Christian
; APPLICANT: Buck, Charles R.
; APPLICANT: Zatloukal, Kurt
; TITLE OF INVENTION: Polypeptides and nucleic acids encoding these and their use for
; TITLE OF INVENTION: prevention, diagnosis or treatment of liver disorders and epith.
; FILE REFERENCE: Oridis Biomed
; CURRENT APPLICATION NUMBER: US/10/528,031
; CURRENT FILING DATE: 2005-03-16
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-528-031-7

Query Match 90.6%; Score 1827.5; DB 9; Length 403;

Best Local Similarity 87.1%; Pred. No. 1.3e-166; Indels 25; Gaps 2;
Matches 351; Conservative 9; Mismatches 18;

QY	1	MSGRVGDLSPRQKEALAK-----PEASTCRSRRPCSGSMWSSES-----	39
Db	1	MSGRVGDLSPRQKEALAKRENVQDVLPALPNDDYFLRLRLRARSFDLQKSEAMLRKHV	60
QY	40	----KRLTTSALWQPEVIQQYLSGCMGCGYLDGCPVWYDIIGPKDAKGLLFSASKQDL	95
Db	61	EPRKQKDIDNIISWQPEVIQQYLSGCMGCGYLDGCPVWYDIIGPKDAKGLLFSASKQDL	120
QY	96	LRTKMECELLLOECAHQTTLGRKRVETITIIYDCEGLGKHLWKPAVEAYGFLCMFEE	155
Db	121	LRTKMECELLLOECAHQTTLGRKRVETITIIYDCEGLGKHLWKPAVEAYGFLCMFEE	180
QY	156	NYPETIKRLVVKAPKLPVAYNLKPFILSEDRKKIMVLGANYKEVLLKHISPDQVPVE	215
Db	181	NYPETIKRLVVKAPKLPVAYNLKPFILSEDRKKIMVLGANYKEVLLKHISPDQVPVE	240
QY	216	YGTGTMIDPGNPKCKSKINYGDIIPRKYVVRDQKQYHSHVQISRGSQHVEYELFPG	275

Db 241 YGWTDPDGNPKCKSKINYGGDI PRKYVVRDQVKKQYEHVSQVISRGSSSHQVEYBILFPG 300
QY 276 CVLRQFMSDGDADVGFGLPKTKMGERQAGMTEVLNQRVNSHLVPEDGTLTCSDDPGI 335
Db 301 CVLRQFMSDGDADVGFGLPKTKMGERQAGMTEVLNQRVNSHLVPEDGTLTCSDDPGI 360
QY 336 YVLRFDNTYSFTHAKKNVFTVEVLLPDKASEEKMQLGAGTPK 378
Db 361 YVLRFDNTYSFTHAKKNVFTVEVLLPDKASEEKMQLGAGTPK 403

RESULT 2

US-11-072-512-2798
; Sequence 2798, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2798
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2798

Query Match 63.0%; Score 1272; DB 11; Length 391;
Best Local Similarity 62.0%; Pred. No. 1.7e-113;
Matches 232; Conservative 58; Mismatches 68; Indels 16; Gaps 2;

QY 20 EASTCRSRPCCGS-----WMSSE--SKRTLTTSLAWQPEVIQQVLSGGM 63
Db 14 ETLTCRNPKTCEGPKSPQPTSCPHALRAWHMFEFRKQDLDNIVTWQPVEIQLYDSGL 73
QY 64 CGYDLGCPVWYDITGPKDAKGLFSASKQDILLRTKMRECELLLOECAHOTTKLGRKYET 123
Db 74 CGDYEGCPVYFNIGSLDPKGLLSASKQDMIRKIRVCELLHECELQTKQGRKIEM 133
QY 124 ITIIVDCGLGLKHLWKPAVEAYGFLCMFEENYPETLKRLEFVVRKAPKLPFVAYNLIRPF 183
Db 134 ALLVFDMEGLSLKHLWKPAVEYVQQFFSILEANYPETLKNLIVIRAPKLPFVAFNLVKSF 193
QY 184 LSEDTRKKI MWJLGANYKEVLLKHISPDQVPVEYGGTMDPDGNPKCKSKINYGGDI PRKY 243
Db 194 MSEETRRKIVILGDNWKQBELTKFISPDQLPVEFGGTMTPDGNPKCLTKINYGGEVPKSY 253
QY 244 YVRDQVKQYEHVSQVISRGSSSHQVEIILFPGCVLRWQFMSDGDADVGFGLPKTKMGERQ 303
Db 254 YLCEQVRLQYEHTRSVGRGSSLOVENEILLFPGCVLRWQFASDGGDIGFGVFLTKMGERQ 313

QY 304 RAGMTEVLNQRVNSHLVPEDGTLTCSDDPGIYVLRFDNTYSFTHAKKNVFTVEVLLPDK 363
Db 314 SAREMTEVLPQRSNAHWPEDGSLTCLQAGVYVLRFDNTYSRMYAKKLSTYVEVLLPDK 373
QY 364 ASEEMKMKQLGAGTP 377
Db 374 ASEETLQSLKAWRP 387

RESULT 3

US-11-072-512-3609
; Sequence 3609, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3609
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3609

Query Match 55.4%; Score 1118; DB 11; Length 360;
Best Local Similarity 66.1%; Pred. No. 7.9e-99;
Matches 197; Conservative 49; Mismatches 52; Indels 0; Gaps 0;

QY 40 KRTLTTSLAWQPEVIQQVLSGGMCGYDLGPKDAKGLFSASKQDILLRTK 99
Db 11 QODLDNIVTWQPPEVIQDYSGGLCYDYGCPVTFNIGSLDPKGLLSASKQDMIRK 70
QY 100 MRECELLLOECAHOTTKLGRKYETITIIYDCEGLGLKHLWKPAVEAYGFLCMFEENYP 159
Db 71 IKVCELLHECELQTKLGRKIEMALMVFDMEGLSLKHLWKPAVEYVQQFFSILEANYPE 130
QY 160 TLKRLFFVVKAPKLPFVAYNLIKPFLSESTRKIMVILGANYKEVLLKHISPDQVPVEYGGT 219
Db 131 TLKNLIVIRAPKLPFVAFNLVKSEETRRKIVILGDNWKQBELTKFISPDQLPVEFGT 190
QY 220 MTDPDGNPKCKSKINYGGDI PRKYVVRDQVKQYEHVSQVISRGSSSHQVEYETILFPGCVLR 279
Db 191 MTDPDGNPKCLTKINYGGEVPASVYLCQVRLQYEHTRSVGRGSSLOVENEILLFPGCVLR 250
QY 280 WQFMSDGDADVGFGLPKTKMGERQAGMTEVLNQRVNSHLVPEDGTLTCSDDPGIYV 337
Db 251 WQFASDGGDIGFGVFLTKMGEQQSAREMTEVLPQRSNAHWPEDGSLTCLQAGVYV 308

US-11-096-568A-28170
; Sequence 28170, Application US/11096568A
; Publication No. US20060048240A1


```

; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28170
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(478)
; OTHER INFORMATION: Ceres Seq. ID no. 2705654
US-11-096-568A-28170

```

```

Query Match      11.0%; Score 221.5; DB 11; Length 478;
Best Local Similarity 26.8%; Pred. No. 6.1e-13;
Matches 70; Conservative 44; Mismatches 116; Indels 31; Gaps 6;

QY 31 SGMWSSESK-----RTLTTSLAWQPPEVIQQYLSGMCYDLDGCPVWYDIIGPKDA 83
Db 16 TNQMSDMLRWKKEFGADTVMEDEFKIDEVLKYPQGHGVKDEGRPVYIERLGQVDS 75

QY 84 KGLLFSASKQDLLRTKMRCE-----LLQECAHQTTKLGKRVETITIIYDCBGLGKLHLW 139
Db 76 TKLMQVTTMDRYVNVHMEFERTFNVKPPACSIAAK---HIDQSTTILDVQGVGLKNFN 132

QY 140 KPAVEAYGEFLCMFEENYPETLKRLLFPVVKAPKLPFVAYNLIKPFLSEDTTRKKIMVLGANY 199
Db 133 KAARDLITRLQKVDGNDNYPETLNRMFIIINAGSGFRMLMNTVKSFLDPKTTAKIHVLGNKY 192

QY 200 KEVLLKHISPDQVPVEYGGTWT-----DPDGNPKCKSKINYGDDIPRKYVYRD 247
Db 193 QSKLLEIIDASELPFLGGSCCTCADNGGCMRSDKGPWNNPDIMKRVNNGDHICSK---RS 249

QY 248 QVKQYEHVSQISRGSSHOVE 268
Db 250 QADNAGENI--ISQGNNSAVE 268

```

```

RESULT 8
US-11-096-568A-28169
; Sequence 28168, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28169
; LENGTH: 517
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(517)
; OTHER INFORMATION: Ceres Seq. ID no. 2705653
US-11-096-568A-28169

```

```

Query Match      11.0%; Score 221.5; DB 11; Length 517;
Best Local Similarity 26.8%; Pred. No. 6.8e-13;
Matches 70; Conservative 44; Mismatches 116; Indels 31; Gaps 6;

QY 31 SGMWSSESK-----RTLTTSLAWQPPEVIQQYLSGMCYDLDGCPVWYDIIGPKDA 83
Db 55 TNQMSDMLRWKKEFGADTVMEDEFKIDEVLKYPQGHGVKDEGRPVYIERLGQVDS 114

```

```

QY 84 KGLLFSASKQDLLRTKMRCE-----LLQECAHQTTKLGKRVETITIIYDCBGLGKLHLW 139
Db 115 TKLMQVTTMDRYVNVHMEFERTFNVKPPACSIAAK---HIDQSTTILDVQGVGLKNFN 171

QY 140 KPAVEAYGEFLCMFEENYPETLKRLLFPVVKAPKLPFVAYNLIKPFLSEDTTRKKIMVLGANY 199
Db 172 KAARDLITRLQKVDGNDNYPETLNRMFIIINAGSGFRMLMNTVKSFLDPKTTAKIHVLGNKY 231

QY 200 KEVLLKHISPDQVPVEYGGTWT-----DPDGNPKCKSKINYGDDIPRKYVYRD 247
Db 232 QSKLLEIIDASELPFLGGSCCTCADNGGCMRSDKGPWNNPDIMKRVNNGDHICSK---RS 288

QY 248 QVKQYEHVSQISRGSSHOVE 268
Db 289 QADNAGENI--ISQGNNSAVE 307

```

RESULT 9

```

US-11-096-568A-28168
; Sequence 28168, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28168
; LENGTH: 558
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(558)
; OTHER INFORMATION: Ceres Seq. ID no. 2705652
US-11-096-568A-28168

```

```

Query Match      11.0%; Score 221.5; DB 11; Length 558;
Best Local Similarity 26.8%; Pred. No. 7.6e-13;
Matches 70; Conservative 44; Mismatches 116; Indels 31; Gaps 6;

```

```

QY 31 SGMWSSESK-----RTLTTSLAWQPPEVIQQYLSGMCYDLDGCPVWYDIIGPKDA 83
Db 96 TNQMSDMLRWKKEFGADTVMEDEFKIDEVLKYPQGHGVKDEGRPVYIERLGQVDS 155

QY 84 KGLLFSASKQDLLRTKMRCE-----LLQECAHQTTKLGKRVETITIIYDCBGLGKLHLW 139
Db 156 TKLMQVTTMDRYVNVHMEFERTFNVKPPACSIAAK---HIDQSTTILDVQGVGLKNFN 212

QY 140 KPAVEAYGEFLCMFEENYPETLKRLLFPVVKAPKLPFVAYNLIKPFLSEDTTRKKIMVLGANY 199
Db 213 KAARDLITRLQKVDGNDNYPETLNRMFIIINAGSGFRMLMNTVKSFLDPKTTAKIHVLGNKY 272

QY 200 KEVLLKHISPDQVPVEYGGTWT-----DPDGNPKCKSKINYGDDIPRKYVYRD 247
Db 273 QSKLLEIIDASELPFLGGSCCTCADNGGCMRSDKGPWNNPDIMKRVNNGDHICSK---RS 329

QY 248 QVKQYEHVSQISRGSSHOVE 268
Db 330 QADNAGENI--ISQGNNSAVE 348

```

RESULT 10

```

US-11-096-568A-32080
; Sequence 32080, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2

```

```
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32080
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(409)
; OTHER INFORMATION: Ceres Seq. ID no. 13592122
US-11-096-568A-32080

Query Match          10.7%; Score 216; DB 11; Length 409;
Best Local Similarity 22.2%; Pred. No. 1.6e-12;
Matches 94; Conservative 58; Mismatches 139; Indels 132; Gaps 17;

QY      5 VGDLSPRQKEAL--AKPEASTCSRSPRCSGSMMS-----S 37
Db      41 VSELKPTQKSLQELKEKLSASSK---ASSMNGVSLGGDDKADVILLKFLRADPKVA 97
QY      38 ESKRTLTTSLAWOPPEVIQY-----LSGG---MCGYDLGCPVWYDIIG---PKD 82
Db      98 DSLRMLEKCLEWREBEFKAELTEEDLGFKDLEGKVAYMRGYDKEGHPVCYNAYGVFKE 157
QY      83 AKGLLFSASKQ--DLLRTKRECELLLQECAGHTTKLGRK-----VETITIIYDCEG 132
Db      158 MYERVFGBDEBKLNKFLRWV-----QVLERGVKMLHPKPGGVNSIIQVTDLKD 205
QY      133 LGLKHLWKPAVEAYGEFLCMFEENYPETLKRFLFVVKAPKLPFVAYNLIKPFLSEDTRKK- 191
Db      206 MPKREL-----RVASNQILSLFQDNYPELVATKIFINVPWFVSIIYMSFPLTQRTSKF 261
QY      192 IMVLGANYKEVLLKHISPDQVPVEYGGTMTDPP---GNPKCKSKINYGDIIPRKYVVRDQ 248
Db      262 VMSKEGNAEATLYKFIREDIPVQYGLSRPTDSQNGPPKPASEFSIKG-----310
QY      249 VKQYEHSHVQISRSSHQVEYIILFPGCVLRWQPMDSGADVGVGIFLTKMGERQORAGEM 308
Db      311 -----GEKVNIIQIEGEGGATITWDIVVGGWDLEYS-----341
QY      309 TEVLNPNQRYNSHLVPED-----GTLTCSDPGIYVLRFDNTYSFIHAKV---N 353
Db      342 AEFVPAEESYAIIVVERPKMKATDEAVCNSTFTVEAGKLILSDVNTLS--RKKKVAAYR 399
QY      354 FTV 356
Db      400 YTV 402

RESULT 11
US-11-096-568A-32079
; Sequence 32079, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32079
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(433)
; OTHER INFORMATION: Ceres Seq. ID no. 13592121
US-11-096-568A-32079

Query Match          10.7%; Score 216; DB 11; Length 433;
```

```
Best Local Similarity 22.2%; Pred. No. 1.8e-12;
Matches 94; Conservative 58; Mismatches 139; Indels 132; Gaps 17;

QY      5 VGDLSPRQKEAL--AKPEASTCSRSPRCSGSMMS-----S 37
Db      65 VSELKPTQKSLQELKEKLSASSK---ASSMNGVSLGGDDKADVILLKFLRADPKVA 121
QY      38 ESKRTLTTSLAWOPPEVIQY-----LSGG---MCGYDLGCPVWYDIIG---PKD 82
Db      122 DSLRMLEKCLEWREBEFKAELTEEDLGFKDLEGKVAYMRGYDKEGHPVCYNAYGVFKE 181
QY      83 AKGLLFSASKQ--DLLRTKRECELLLQECAGHTTKLGRK-----VETITIIYDCEG 132
Db      182 MYERVFGBDEBKLNKFLRWV-----QVLERGVKMLHPKPGGVNSIIQVTDLKD 229
QY      133 LGLKHLWKPAVEAYGEFLCMFEENYPETLKRFLFVVKAPKLPFVAYNLIKPFLSEDTRKK- 191
Db      230 MPKREL-----RVASNQILSLFQDNYPELVATKIFINVPWFVSIIYMSFPLTQRTSKF 285
QY      192 IMVLGANYKEVLLKHISPDQVPVEYGGTMTDPP---GNPKCKSKINYGDIIPRKYVVRDQ 248
Db      286 VMSKEGNAEATLYKFIREDIPVQYGLSRPTDSQNGPPKPASEFSIKG-----334
QY      249 VKQYEHSHVQISRSSHQVEYIILFPGCVLRWQPMDSGADVGVGIFLTKMGERQORAGEM 308
Db      335 -----GEKVNIIQIEGEGGATITWDIVVGGWDLEYS-----365
QY      309 TEVLNPNQRYNSHLVPED-----GTLTCSDPGIYVLRFDNTYSFIHAKV---N 353
Db      366 AEFVPAEESYAIIVVERPKMKATDEAVCNSTFTVEAGKLILSDVNTLS--RKKKVAAYR 423
QY      354 FTV 356
Db      424 YTV 426

RESULT 12
US-11-096-568A-27441
; Sequence 27441, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27441
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(511)
; OTHER INFORMATION: Ceres Seq. ID no. 1815699
US-11-096-568A-27441

Query Match          10.6%; Score 214; DB 11; Length 511;
Best Local Similarity 28.7%; Pred. No. 3.5e-12;
Matches 58; Conservative 35; Mismatches 95; Indels 14; Gaps 3;

QY      34 MWSSE-----SKRTLTSLAWOPPEVIQYQLSGGCMGYDLGCPVWYDIIGPKDAKGL 86
Db      57 MWSDMIQWRKDFGADTTIEDFDEIDVEMKHPYQGVHGVDEKGRPVIERLGGIDANKL 116
QY      87 LFSASKQDLLRTKRECE-----LILQECAGHTTKLGRKVFETITIIYDCEGLGLKHLWKPA 142
Db      117 LQVTTMDRYKYHYKREKTEFKVKFPSCSVAANK---HIDOSTILDVQGVGLKNFSKA 173
QY      143 VEAYGEFLCMFEENYPETLKRFLFVVKAPKLPFVAYNLIKPFLSEDTRKKIMVLGANYKEV 202
Db      174 RELQRLCKIDNENYPETLNRMFIIINAGSFRLLWSTVKSPLDPKTTAKIHVLGNKYHSK 233
```

QY 203 LLKHISPDQVPVEYGGTMDPD 224
||:|::|:|:|
Db 234 LLEVIDASELPFEFGGACTCED 255

```

RESULT 13
US-11-096-568A-27440
; Sequence 27440, Application US/11096568A
; Publication NO. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA F
; TITLE OF INVENTION: Theryb
; FILE REFERENCE: 2750-1532PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27440
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(551)
; OTHER INFORMATION: Ceres Seq. ID NO. 1815698
US-11-096-568A-27440

```

Query Match	10.6%	Score 214;	DB 11;	Length 551;
Best Local Similarity	28.7%;	Pred. No. 3.9e-11;		
Matches 58;	Conservative 35;	Mismatches 95;	Indels 14;	Gaps 3;
Qy	34	MWSE-----SKRTLTTSIAWQPEVIOQYLSGSGCMGYDLDCGPVWYDIIGPKDAKGL	86	
Db	97	MWSDMIQWRKDFGADTIIEDPDEEIDVEMKHYPQYGVHDKEGRPVVIERLGQIDANKL	156	
Qy	87	LFSASKQDLRTKWRCE-----LLQECANHTTKLGRKVETITIIYDCEGLGLKHLWKPA	142	
Db	157	LQVTTDRYKYVHYKFEKTFKVKFPSPCSVAANK---HIDQSTTILDVQGVGLKNFSKSA	213	
Qy	143	VEAYGEFLCFMEENYPETLKRFLVVKAPKLPFVAYNLIKPFLSEDTKKIMVLGANYKEV	202	
Db	214	RELLQRLCKTDNENYPETLNMFIINAGSGPRLIWSVKSFLDPKTTAKIHVLGNKYHSK	273	
Qy	203	LLKHISPDQVFVEYGGTMTDDP	224	
Db	274	LLEVIDASELPFEFFGACTCD	295	

```

RESULT 14
US-11-096-568A-27439
; Sequence 27439, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA F
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592APUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27439
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(560)
; OTHER INFORMATION: Ceres Seq. ID no. 1815697
US-11-096-568A-27439

```

Query Match 10.6%; Score 214; DB 11; Length 560;
Best Local Similarity 28.7%; Pred. No. 4e-12;

[illegible]

```

RESULT 15
US-11-096-568A-9415
; Sequence 9415, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fr
; TITLE OF INVENTION: Theryb
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 9415
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(394)
; OTHER INFORMATION: Ceres Seq. ID no. 12460984
US-11-096-568A-9415

```

Query Match	10.6%	Score 213;	DB 11;	Length 394;
Best Local Similarity	24.9%;	Pred. No. 3e-12;		
Matches	80;	Conservative 46;	Mismatches 147;	Indels 48; Gaps 10;
QY	:	:	:	:
Db	:	:	:	:
QY	35	WSSB--SKRTLTTS	SLAMQPPEV	IQYLSGGWCGYDLGCPWYDIIG---PKDAKGLLFS 89
Db	95	WRASFRAVDLDELGF	DLLEGIVAY---	MHGWRDGHPCVYNAYGVFKDRMDYRVFG 150
QY	90	ASKQDLARTKMRCE	LLQBCAHQTT	KLGRKVKETITIIYDCEGLGLKHLWKPAVEAYGEF 149
Db	151	DG--DRLSRFLRW	RQVIMERGVALQ	LRPGGVNAIIQVTDLKDMPKREL-----RAASNQI 204
QY	150	LCMEENYPETLKL	FVVPKAPKLP	FPVAYNLIKPLSEDTKRIKIVL--GANYKEVLLKHIS 208
Db	205	LSLFQDNYPEMVAR	KVFVNVPWYF	SVLSFLSMISPLFTEKTSKFVIAREGNVAETLFKFIR 264
QY	209	PDQVPVBYGGTMD	DPD---GNPKCKS	KINYGDDIPRKYVYVRDQVKQVQHSVQISRGSSH 265
Db	265	PELVVPVQVGGLS	RASELENGEP	PKPASFTTNG-----GKVV 300
QY	266	QVEYEILFP	PGCVLRQFMS	DGADVGFGI--FLUKTKGERQRAGEMTEVLPNQRYNSHLVPE 324
Db	301	FLEIDGIEAGATI	TDWLIVVGWE	LEYGAEYYVPASIEGGYTLVCVETKVKVPAAADE----PV 356
QY	325	DGTLTCS	DPGIVYLR	FDNTYS 345
Db	357	HNAFTAKE	PGKQVLS	IDNSGS 377

Search completed: May 2, 2006, 18:53:58
Job time : 28 secs

PEAST CRSRRPCSGSMWSSSKRTLTTSLA

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 18:34:32 ; Search time 40 Seconds
(without alignments)
909.248 Million cell updates/sec

Title: US-10-696-699A-2
Perfect score: 2018
Sequence: 1 MGRVGDLSPRQKEALAKPE.....LLPDKASEEKKQLGAGTPK 378

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1827.5	90.6	403	2 JC7708	tocopherol associa
2	661.5	32.8	343	2 A53057	retinal-binding pr
3	408	20.2	383	2 T23057	hypothetical prote
4	388.5	19.3	470	2 S28303	hypothetical prote
5	388.5	19.3	743	2 B88551	protein T23G5.2 li
6	334	16.6	1254	2 T24897	hypothetical prote
7	229.5	11.4	490	2 A96745	probable cytosolic
8	229	11.3	457	2 C86329	hypothetical prote
9	228	11.3	286	2 T38768	probable sec14 cyt
10	221.5	11.0	558	2 C84561	hypothetical prote
11	218	10.8	582	2 F84539	hypothetical prote
12	217.5	10.8	371	2 B84602	hypothetical prote
13	216	10.7	409	2 T46063	hypothetical prote
14	214	10.6	560	2 T05278	hypothetical prote
15	210	10.4	301	2 S57923	SEC14 protein - ye
16	210	10.4	310	2 S37916	SEC14 protein homo
17	207	10.3	558	2 G85430	hypothetical prote
18	206.5	10.2	531	2 C84602	hypothetical prote
19	204.5	10.1	396	2 T33386	hypothetical prote
20	203	10.1	723	2 H85092	hypothetical prote
21	199.5	9.9	550	2 T08566	hypothetical prote
22	197	9.8	304	2 A30106	SEC14 protein - ye
23	188	9.3	341	2 G89500	protein T27A10.7
24	183.5	9.1	640	2 B96784	hypothetical prote
25	181	9.0	278	1 A47404	alpha-tocopherol t
26	181	9.0	278	2 S54352	alpha-tocopherol t
27	179	8.9	377	2 T21170	hypothetical prote
28	179	8.9	453	2 T21528	hypothetical prote
29	176.5	8.7	497	2 S43745	phosphatidylinosit

SEC14 protein - ye
hypothetical prote
hypothetical prote
probable cytosolic
hypothetical prote
phosphatidylinosit
hypothetical prote
F12K8.13 protein -
hypothetical prote
hypothetical prote
polyphosphoinosit
SEC14 protein homo
hypothetical prote
protein F14J16.8
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

JC7708
tocopherol associated protein - human
N;Alternate names: alpha-tocopherol associated protein
C;Species: Homo sapiens (man)
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Jul-2004
C;Accession: JC7708
R;Yamauchi, J.; Iwamoto, T.; Kida, S.; Masushige, S.; Yamada, K.; Esashi, T.
Biochem. Biophys. Res. Commun. 285, 295-299, 2001
A;Title: Tocopherol-associated protein is a ligand-dependent transcriptional activator.
A;Reference number: JC7708; MUID:21338208; PMID:11444841
A;Contents: Liver
A;Accession: JC7708
A;Molecule type: mRNA
A;Residues: 1-403 <YAM>
A;Cross-references: UNIPROT:O76054; UNIPARC:UPI000017CBA2; GB:AL096881
C;Comment: This protein has the following properties: (i) alpha-tocopherol specific bin
driptional activation involved in gene regulation in vivo in mammalian cells.
C;Genetics:
A;Gene: tap

Query Match		90.6%;	Score 1827.5;	DB 2;	Length 403;
Best Local Similarity		87.1%;	Pred. No. 2.4e-147;		
Matches 351;		Conservative	9;	Mismatches 18;	Indels 25; Gaps 2;
Qy	1	MGRVGDLSPRQKEALAK	-----PEASTCRSRPCSGSMWSSS-----	39	
Db	1	MGRVGDLSPRQKEALAKFRENVDVLPALPNPDHFLLEWLARSFDLQKSEAMLRKHV	60		
Qy	40	-----KRTLTTSLAWQPEVITQOYLSGCMCGYDLDGCPVWYDIIGPKDAKGLLPASAKQDL	95		
Db	61	EPRKQKDIDNIISWQPEVITQOYLSGCMCGYDLDGCPVWYDIIGPLDAKGLLPASAKQDL	120		
Qy	96	LRTKMRCELLQECACHQTTKLRKVKETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE	155		
Db	121	LRTKMRCELLQECACHQTTKLRKVKETITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEE	180		
Qy	156	NYPETLKRFLVVKAPKLPVAYNLIKPFLSEDTTRKKIMVLGANYKEVLLKXHSIDPDQVPE	215		
Db	181	NYPETLKRFLVVKAPKLPVAYNLIKPFLSEDTTRKKIMVLGANYKEVLLKXHSIDPDQVPE	240		
Qy	216	YGGTMTDPDGNPKCKSKINYGDIIPKYYVRDQKQYEHVSQISRGSSHQVEYELLPPG	275		
Db	241	YGGTMTDPDGNPKCKSKINYGDIIPKYYVRDQKQYEHVSQISRGSSHQVEYELLPPG	300		
Qy	276	CVLRWQFMSDQADVFGIFLTKMGERQORAGETEVLPNQRYNSHLVPEQDTLLFCSPDGI	335		
Db	301	CVLRWQFMSDQADVFGIFLTKMGERQORAGETEVLPNQRYNSHLVPEQDTLLFCSPDGI	360		
Qy	336	YVLRFDNTYSFIHAKKNVFTVEVLLPDKASEEKKQLGAGTPK	378		
Db	361	YVLRFDNTYSFIHAKKNVFTVEVLLPDKASEEKKQLGAGTPK	403		

Qy	49	WQPEVVIQOYLSGGCGYDLDCGPVWYDIIGPKDAKGLLFSASKQDLLRTKRVCELLIQ	108
Db	71	WTPPECLEKICYGGLG-DTEGRFILMSLGNVDVEGLRSVASLDYIFKSLAAIEKGMK	129
Qy	109	ECAHQTTKLGRKVEITIIYDCEGLGLKHL-WKPAVEAYGEFLCMFEENYPETLKRFLVV	167
Db	130	LCEEAKESGRPEQNTLVDFDENITSAHFSCKQFASSTTIVSLUFODHYPLFLRKILLI	189
Qy	168	KAPKLPVAYNLIKPFLESD-TRKKIMVLGANYKEVLLKHISPDQVPVVEYGGTMTDPDGN	226
Db	190	RAPEMARIAYSITAILQDPIRLVEMPSESDKWSLAQIVNLDAWPMYWGGNLVE-NGD	248
Qy	227	PKCKSKINYGDIIPKRYVVRDQV--QQVEHSVQISRGSSHQVEYELIPGCVLVRQWQFMS	284
Db	249	PKCPSRIKRYGGGAVDESVPDPKKAMADYDQLTTVYAGDKHLIQIKVRPSRI-SWTYMT	307
Qy	285	DGADVGGFGLTKMGERORACEMTEVLNQYNSHLVPEDEGTLCSDPGIYVLRFDNTY	344
Db	308	DEDDLGFEIHY-DKTGSCDKLTEMETVPIYIRLECTNVEITGHLDVTDVGNVLEFDNY	366
Qy	345	SPIHAKKNFTVEV	358
Db	367	SWFSAKQLRYNIEI	380
<p>RESULT 4</p> <p>S28303</p> <p>hypothetical protein T23G5.2 - <i>Caenorhabditis elegans</i></p> <p>C;Species: <i>Caenorhabditis elegans</i></p> <p>C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004</p> <p>C;Accession: S28303</p> <p>R;Berk, M.</p> <p>submitted to the EMBL Data Library, December 1992</p> <p>A;Reference number: S28296</p> <p>A;Accession: S28303</p> <p>A;Molecule type: DNA</p> <p>A;Residues: 1-470 <BER></p> <p>A;Cross-references: UNIPROT:Q03606; UNIPARC:UPI000017BBE0; EMBL:Z19158</p> <p>C;Genetics:</p> <p>A;Introns: 209/3; 322/3; 389/3; 439/2</p> <p>F;32-240/Domain: cellular retinaldehyde-binding protein homology <CRB></p>			
<p>Query Match 19.3%; Score 388.5; DB 2; Length 470;</p> <p>Best Local Similarity 28.2%; Pred. No. 5.3e-25;</p> <p>Matches 106; Conservative 59; Mismatches 136; Indels 75; Gaps 10;</p>			
Qy	49	WQPEVVIQOYLSGGCGYDLDCGPVWYDIIGPKDAKGLLFSASKQDLLRTKRVCELLIQ	108
Db	68	WTRPTVIKQYFPGCWHSNDKAGRPYILRFGQLDTKGLRSCGVENLVKLTLISICEDGLQ	127
Qy	109	ECAHQTTKLGRKVEITIIYDCEGLGLKHLWKPAVEAYGEFLCMFEENYPETLKRFLVVK	168
Db	128	RAEAATRKLTGPSLSSWSLVVDLDGLSMRHLWRPGVQCLLKIIIEVANIPEINMGVLVVR	187
Qy	169	APKLPVAYNLIKPFLESDTRKKIMV---LGANYKEVLLKHISPDQVPVVEYGGTMTDPDG	225
Db	188	APRVFVPLWTLSPFIDEXTRKFMVSGSGGDLKEELAKHIEKEKIPDFLGSSC-----	242
Qy	226	NPCKSKINYGDIIPKRYVVRDQVQOYE-----HSVQISRGSSHQVEYELIPF---	274
Db	243	---LTTNCGILGGHPVKSMYL--PVBEQEGASSSEDLPHSTYTSTATWRGYPVEWPIET	297
Qy	275	-GCVLRWQFMSDQADVGGIFLTKMGE---RQRAGEMTEVL-----PNOR	316
Db	298	AGCVLTWDFDLVKNDCEFLSYFSTEKIEQPAVRDGAQSPTTLNPFVEMVSAIIGGASHQH	357
Qy	317	YNSHLVE-----DGLTLCSDPGIYVLR-----DNTY	344
Db	358	PDLOCAPELKIQTPLRLKEKAVVFGDSDMQGSHYCSRAGTYIMQWRVPETAAGHSSTF	417
Qy	345	SP-IHAKKNFTVEVL	359
Db	418	DFGSHKCRLIYYEIL	433

A;Map position: 2
A;Introns: 63/1; 158/2; 200/3; 286/2; 404/3; 437/1; 539/3; 616/1; 688/3; 732/3; 778/3; 816/1; 888/2; 958/3; 1028/2; 1108/3; 1188/2; 1268/3; 1348/2; 1428/3; 1508/2; 1588/3; 1668/2; 1748/3; 1828/2; 1908/3; 1988/2; 2068/3; 2148/2; 2228/3; 2308/2; 2388/3; 2468/2; 2548/3; 2628/2; 2708/3; 2788/2; 2868/3; 2948/2; 3028/3; 3108/2; 3188/3; 3268/2; 3348/3; 3428/2; 3508/3; 3588/2; 3668/3; 3748/2; 3828/3; 3908/2; 3988/3; 4068/2; 4148/3; 4228/2; 4308/3; 4388/2; 4468/3; 4548/2; 4628/3; 4708/2; 4788/3; 4868/2; 4948/3; 5028/2; 5108/3; 5188/2; 5268/3; 5348/2; 5428/3; 5508/2; 5588/3; 5668/2; 5748/3; 5828/2; 5908/3; 5988/2; 6068/3; 6148/2; 6228/3; 6308/2; 6388/3; 6468/2; 6548/3; 6628/2; 6708/3; 6788/2; 6868/3; 6948/2; 7028/3; 7108/2; 7188/3; 7268/2; 7348/3; 7428/2; 7508/3; 7588/2; 7668/3; 7748/2; 7828/3; 7908/2; 7988/3; 8068/2; 8148/3; 8228/2; 8308/3; 8388/2; 8468/3; 8548/2; 8628/3; 8708/2; 8788/3; 8868/2; 8948/3; 9028/2; 9108/3; 9188/2; 9268/3; 9348/2; 9428/3; 9508/2; 9588/3; 9668/2; 9748/3; 9828/2; 9908/3; 9988/2; 10068/3; 10148/2; 10228/3; 10308/2; 10388/3; 10468/2; 10548/3; 10628/2; 10708/3; 10788/2; 10868/3; 10948/2; 11028/3; 11108/2; 11188/3; 11268/2; 11348/3; 11428/2; 11508/3; 11588/2; 11668/3; 11748/2; 11828/3; 11908/2; 11988/3; 12068/2; 12148/3; 12228/2; 12308/3; 12388/2; 12468/3; 12548/2; 12628/3; 12708/2; 12788/3; 12868/2; 12948/3; 13028/2; 13108/3; 13188/2; 13268/3; 13348/2; 13428/3; 13508/2; 13588/3; 13668/2; 13748/3; 13828/2; 13908/3; 13988/2; 14068/3; 14148/2; 14228/3; 14308/2; 14388/3; 14468/2; 14548/3; 14628/2; 14708/3; 14788/2; 14868/3; 14948/2; 15028/3; 15108/2; 15188/3; 15268/2; 15348/3; 15428/2; 15508/3; 15588/2; 15668/3; 15748/2; 15828/3; 15908/2; 15988/3; 16068/2; 16148/3; 16228/2; 16308/3; 16388/2; 16468/3; 16548/2; 16628/3; 16708/2; 16788/3; 16868/2; 16948/3; 17028/2; 17108/3; 17188/2; 17268/3; 17348/2; 17428/3; 17508/2; 17588/3; 17668/2; 17748/3; 17828/2; 17908/3; 17988/2; 18068/3; 18148/2; 18228/3; 18308/2; 18388/3; 18468/2; 18548/3; 18628/2; 18708/3; 18788/2; 18868/3; 18948/2; 19028/3; 19108/2; 19188/3; 19268/2; 19348/3; 19428/2; 19508/3; 19588/2; 19668/3; 19748/2; 19828/3; 19908/2; 19988/3; 20068/2; 20148/3; 20228/2; 20308/3; 20388/2; 20468/3; 20548/2; 20628/3; 20708/2; 20788/3; 20868/2; 20948/3; 21028/2; 21108/3; 21188/2; 21268/3; 21348/2; 21428/3; 21508/2; 21588/3; 21668/2; 21748/3; 21828/2; 21908/3; 21988/2; 22068/3; 22148/2; 22228/3; 22308/2; 22388/3; 22468/2; 22548/3; 22628/2; 22708/3; 22788/2; 22868/3; 22948/2; 23028/3; 23108/2; 23188/3; 23268/2; 23348/3; 23428/2; 23508/3; 23588/2; 23668/3; 23748/2; 23828/3; 23908/2; 23988/3; 24068/2; 24148/3; 24228/2; 24308/3; 24388/2; 24468/3; 24548/2; 24628/3; 24708/2; 24788/3; 24868/2; 24948/3; 25028/2; 25108/3; 25188/2; 25268/3; 25348/2; 25428/3; 25508/2; 25588/3; 25668/2; 25748/3; 25828/2; 25908/3; 25988/2; 26068/3; 26148/2; 26228/3; 26308/2; 26388/3; 26468/2; 26548/3; 26628/2; 26708/3; 26788/2; 26868/3; 26948/2; 27028/3; 27108/2; 27188/3; 27268/2; 27348/3; 27428/2; 27508/3; 27588/2; 27668/3; 27748/2; 27828/3; 27908/2; 27988/3; 28068/2; 28148/3; 28228/2; 28308/3; 28388/2; 28468/3; 28548/2; 28628/3; 28708/2; 28788/3; 28868/2; 28948/3; 29028/2; 29108/3; 29188/2; 29268/3; 29348/2; 29428/3; 29508/2; 29588/3; 29668/2; 29748/3; 29828/2; 29908/3; 29988/2; 30068/3; 30148/2; 30228/3; 30308/2; 30388/3; 30468/2; 30548/3; 30628/2; 30708/3; 30788/2; 30868/3; 30948/2; 31028/3; 31108/2; 31188/3; 31268/2; 31348/3; 31428/2; 31508/3; 31588/2; 31668/3; 31748/2; 31828/3; 31908/2; 31988/3; 32068/2; 32148/3; 32228/2; 32308/3; 32388/2; 32468/3; 32548/2; 32628/3; 32708/2; 32788/3; 32868/2; 32948/3; 33028/2; 33108/3; 33188/2; 33268/3; 33348/2; 33428/3; 33508/2; 33588/3; 33668/2; 33748/3; 33828/2; 33908/3; 33988/2; 34068/3; 34148/2; 34228/3; 34308/2; 34388/3; 34468/2; 34548/3; 34628/2; 34708/3; 34788/2; 34868/3; 34948/2; 35028/3; 35108/2; 35188/3; 35268/2; 35348/3; 35428/2; 35508/3; 35588/2; 35668/3; 35748/2; 35828/3; 35908/2; 35988/3; 36068/2; 36148/3; 36228/2; 36308/3; 36388/2; 36468/3; 36548/2; 36628/3; 36708/2; 36788/3; 36868/2; 36948/3; 37028/2; 37108/3; 37188/2; 37268/3; 37348/2; 37428/3; 37508/2; 37588/3; 37668/2; 37748/3; 37828/2; 37908/3; 37988/2; 38068/3; 38148/2; 38228/3; 38308/2; 38388/3; 38468/2; 38548/3; 38628/2; 38708/3; 38788/2; 38868/3; 38948/2; 39028/3; 39108/2; 39188/3; 39268/2; 39348/3; 39428/2; 39508/3; 39588/2; 39668/3; 39748/2; 39828/3; 39908/2; 39988/3; 40068/2; 40148/3; 40228/2; 40308/3; 40388/2; 40468/3; 40548/2; 40628/3; 40708/2; 40788/3; 40868/2; 40948/3; 41028/2; 41108/3; 41188/2; 41268/3; 41348/2; 41428/3; 41508/2; 41588/3; 41668/2; 41748/3; 41828/2; 41908/3; 41988/2; 42068/3; 42148/2; 42228/3; 42308/2; 42388/3; 42468/2; 42548/3; 42628/2; 42708/3; 42788/2; 42868/3; 42948/2; 43028/3; 43108/2; 43188/3; 43268/2; 43348/3; 43428/2; 43508/3; 43588/2; 43668/3; 43748/2; 43828/3; 43908/2; 43988/3; 44068/2; 44148/3; 44228/2; 44308/3; 44388/2; 44468/3; 44548/2; 44628/3; 44708/2; 44788/3; 44868/2; 44948/3; 45028/2; 45108/3; 45188/2; 45268/3; 45348/2; 45428/3; 45508/2; 45588/3; 45668/2; 45748/3; 45828/2; 45908/3; 45988/2; 46068/3; 46148/2; 46228/3; 46308/2; 46388/3; 46468/2; 46548/3; 46628/2; 46708/3; 46788/2; 46868/3; 46948/2; 47028/3; 47108/2; 47188/3; 47268/2; 47348/3; 47428/2; 47508/3; 47588/2; 47668/3; 47748/2; 47828/3; 47908/2; 47988/3; 48068/2; 48148/3; 48228/2; 48308/3; 48388/2; 48468/3; 48548/2; 48628/3; 48708/2; 48788/3; 48868/2; 48948/3; 49028/2; 49108/3; 49188/2; 49268/3; 49348/2; 49428/3; 49508/2; 49588/3; 49668/2; 49748/3; 49828/2; 49908/3; 49988/2; 50068/3; 50148/2; 50228/3; 50308/2; 50388/3; 50468/2; 50548/3; 50628/2; 50708/3; 50788/2; 50868/3; 50948/2; 51028/3; 51108/2; 51188/3; 51268/2; 51348/3; 51428/2; 51508/3; 51588/2; 51668/3; 51748/2; 51828/3; 51908/2; 51988/3; 52068/2; 52148/3; 52228/2; 52308/3; 52388/2; 52468/3; 52548/2; 52628/3; 52708/2; 52788/3; 52868/2; 52948/3; 53028/2; 53108/3; 53188/2; 53268/3; 53348/2; 53428/3; 53508/2; 53588/3; 53668/2; 53748/3; 53828/2; 53908/3; 53988/2; 54068/3; 54148/2; 54228/3; 54308/2; 54388/3; 54468/2; 54548/3; 54628/2; 54708/3; 54788/2; 54868/3; 54948/2; 55028/3; 55108/2; 55188/3; 55268/2; 55348/3; 55428/2; 55508/3; 55588/2; 55668/3; 55748/2; 55828/3; 55908/2; 55988/3; 56068/2; 56148/3; 56228/2; 56308/3; 56388/2; 56468/3; 56548/2; 56628/3; 56708/2; 56788/3; 56868/2; 56948/3; 57028/2; 57108/3; 57188/2; 57268/3; 57348/2; 57428/3; 57508/2; 57588/3; 57668/2; 57748/3; 57828/2; 57908/3; 57988/2; 58068/3; 58148/2; 58228/3; 58308/2; 58388/3; 58468/2; 58548/3; 58628/2; 58708/3; 58788/2; 58868/3; 58948/2; 59028/3; 59108/2; 59188/3; 59268/2; 59348/3; 59428/2; 59508/3; 59588/2; 59668/3; 59748/2; 59828/3; 59908/2; 59988/3; 60068/2; 60148/3; 60228/2; 60308/3; 60388/2; 60468/3; 60548/2; 60628/3; 60708/2; 60788/3; 60868/2; 60948/3; 61028/2; 61108/3; 61188/2; 61268/3; 61348/2; 61428/3; 61508/2; 61588/3; 61668/2; 61748/3; 61828/2; 61908/3; 61988/2; 62068/3; 62148/2; 62228/3; 62308/2; 62388/3; 62468/2; 62548/3; 62628/2; 62708/3; 62788/2; 62868/3; 62948/2; 63028/3; 63108/2; 63188/3; 63268/2; 63348/3; 63428/2; 63508/3; 63588/2; 63668/3; 63748/2; 63828/3; 63908/2; 63988/3; 64068/2; 64148/3; 64228/2; 64308/3; 64388/2; 64468/3; 64548/2; 64628/3; 64708/2; 64788/3; 64868/2; 64948/3; 65028/2; 65108/3; 65188/2; 65268/3; 65348/2; 65428/3; 65508/2; 65588/3; 65668/2; 65748/3; 65828/2; 65908/3; 65988/2; 66068/3; 66148/2; 66228/3; 66308/2; 66388/3; 66468/2; 66548/3; 66628/2; 66708/3; 66788/2; 66868/3; 66948/2; 67028/3; 67108/2; 67188/3; 67268/2; 67348/3; 67428/2; 67508/3; 67588/2; 67668/3; 67748/2; 67828/3; 67908/2; 67988/3; 68068/2; 68148/3; 68228/2; 68308/3; 68388/2; 68468/3; 68548/2; 68628/3; 68708/2; 68788/3; 68868/2; 68948/3; 69028/2; 69108/3; 69188/2; 69268/3; 69348/2; 69428/3; 69508/2; 69588/3; 69668/2; 69748/3; 69828/2; 69908/3; 69988/2; 70068/3; 70148/2; 70228/3; 70308/2; 70388/3; 70468/2; 70548/3; 70628/2; 70708/3; 70788/2; 70868/3; 70948/2; 71028/3; 71108/2; 71188/3; 71268/2; 71348/3; 71428/2; 71508/3; 71588/2; 71668/3; 71748/2; 71828/3; 71908/2; 71988/3; 72068/2; 72148/3; 72228/2; 72308/3; 72388/2; 72468/3; 72548/2; 72628/3; 72708/2; 72788/3; 72868/2; 72948/3; 73028/2; 73108/3; 73188/2; 73268/3; 73348/2; 73428/3; 73508/2; 73588/3; 73668/2; 73748/3; 73828/2; 73908/3; 73988/2; 74068/3; 74148/2; 74228/3; 74308/2; 74388/3; 74468/2; 74548/3; 74628/2; 74708/3; 74788/2; 74868/3; 74948/2; 75028/3; 75108/2; 75188/3; 75268/2; 75348/3; 75428/2; 75508/3; 75588/2; 75668/3; 75748/2; 75828/3; 75908/2; 75988/3; 76068/2; 76148/3; 76228/2; 76308/3; 76388/2; 76468/3; 76548/2; 76628/3; 76708/2; 76788/3; 76868/2; 76948/3; 77028/2; 77108/3; 77188/2; 77268/3; 77348/2; 77428/3; 77508/2; 77588/3; 77668/2; 77748/3; 77828/2; 77908/3; 77988/2; 78068/3; 78148/2; 78228/3; 78308/2; 78388/3; 78468/2; 78548/3; 78628/2; 78708/3; 78788/2; 78868/3; 78948/2; 79028/3; 79108/2; 79188/3; 79268/2; 79348/3; 79428/2; 79508/3; 79588/2; 79668/3; 79748/2; 79828/3; 79908/2; 79988/3; 80068/2; 80148/3; 80228/2; 80308/3; 80388/2; 80468/3; 80548/2; 80628/3; 80708/2; 80788/3; 80868/2; 80948/3; 81028/2; 81108/3; 81188/2; 81268/3; 81348/2; 81428/3; 81508/2; 81588/3; 81668/2; 81748/3; 81828/2; 81908/3; 81988/2; 82068/3; 82148/2; 82228/3; 82308/2; 82388/3; 82468/2; 82548/3; 82628/2; 82708/3; 82788/2; 82868/3; 82948/2; 83028/3; 83108/2; 83188/3; 83268/2; 83348/3; 83428/2; 83508/3; 83588/2; 83668/3; 83748/2; 83828/3; 83908/2; 83988/3; 84068/2; 84148/3; 84228/2; 84308/3; 84388/2; 84468/3; 84548/2; 84628/3; 84708/2; 84788/3; 84868/2; 84948/3; 85028/2; 85108/3; 85188/2; 85268/3; 85348/2; 85428/3; 85508/2; 85588/3; 85668/2; 85748/3; 85828/2; 85908/3; 85988/2; 86068/3; 86148/2; 86228/3; 86308/2; 86388/3; 86468/2; 86548/3; 86628/2; 86708/3; 86788/2; 86868/3; 86948/2; 87028/3; 87108/2; 87188/3; 87268/2; 87348/3; 87428/2; 87508/3; 87588/2; 87668/3; 87748/2; 87828/3; 87908/2; 87988/3; 88068/2; 88148/3; 88228/2; 88308/3; 88388/2; 88468/3; 88548/2; 88628/3; 88708/2; 88788/3; 88868/2; 88948/3; 89028/2; 89108/3; 89188/2; 89268/3; 89348/2; 89428/3; 89508/2; 89588/3; 89668/2; 89748/3; 89828/2; 89908/3; 89988/2; 90068/3; 90148/2; 90228/3; 90308/2; 90388/3; 90468/2; 90548/3; 90628/2; 90708/3; 90788/2; 90868/3; 90948/2; 91028/3; 91108/2; 91188/3; 91268/2; 91348/3; 91428/2; 91508/3; 91588/2; 91668/3; 91748/2; 91828/3; 91908/2; 91988/3; 92068/2; 92148/3; 92228/2; 92308/3; 92388/2; 92468/3; 92548/2; 92628/3; 92708/2; 92788/3; 92868/2; 92948/3; 93028/2; 93108/3; 93188/2; 93268/3; 93348/2; 93428/3; 93508/2; 93588/3; 93668/2; 93748/3; 93828/2; 93908/3; 93988/2; 94068/3; 94148/2; 94228/3; 94308/2; 94388/3; 94468/2; 94548/3; 94628/2; 94708/3; 94788/2; 94868/3; 94948/2; 95028/3; 95108/2; 95188/3; 95268/2; 95348/3; 95428/2; 95508/3; 95588/2; 95668/3; 95748/2; 95828/3; 95908/2; 95988/3; 96068/2; 96148/3; 96228/2; 96308/3; 96388/2; 96468/3; 96548/2; 96628/3; 96708/2; 96788/3; 96868/2; 96948/3; 97028/2; 97108/3; 97188/2; 97268/3; 97348/2; 97428/3; 97508/2; 97588/3; 97668/2; 97748/3; 97828/2; 97908/3; 97988/2; 98068/3; 98148/2; 98228/3; 98308/2; 98388/3; 98468/2; 98548/3; 98628/2; 98708/3; 98788/2; 98868/3; 98948/2; 99028/3; 99108/2; 99188/3; 99268/2; 99348/3; 99428/2; 99508/3; 99588/2; 99668/3; 99748/2; 99828/3; 99908/2; 99988/3; 100068/2; 100148/3; 100228/2; 100308/3; 100388/2; 100468/3; 100548/2; 100628/3; 100708/2; 100788/3; 100868/2; 100948/3; 101028/2; 101108/3; 101188/2; 101268/3; 101348/2; 101428/3; 101508/2; 101588/3; 101668/2; 101748/3; 101828/2; 101908/3; 101988/2; 102068/3; 102148/2; 102228/3; 102308/2; 102388/3; 102468/2; 102548/3; 102628/2; 102708/3; 102788/2; 102868/3; 102948/2; 103028/3; 103108/2; 103188/3; 103268/2; 103348/3; 103428/2; 103508/3; 103588/2; 103668/3; 103748/2; 103828/3; 103908/2; 103988/3; 104068/2; 104148/3; 104228/2; 104308/3; 104388/2; 104468/3; 104548/2; 104628/3; 104708/2; 104788/3; 104868/2; 104948/3; 105028/2; 105108/3; 105188/2; 105268/3; 105348/2; 105428/3; 105508/2; 105588/3; 105668/2; 105748/3; 105828/2; 105908/3; 105988/2; 106068/3; 106148/2; 106228/3; 106308/2; 106388/3; 106468/2; 106548/3; 106628/2; 106708/3; 106788/2; 106868/3; 106948/2; 107028/3; 107108/2; 107188/3; 107268/2; 107348/3; 107428/2; 107508/3; 107588/2; 107668/3; 107748/2; 107828/3; 107908/2; 107988/3; 108068/2; 108148/3; 108228/2; 108308/3; 108388/2; 108468/3; 108548/2; 108628/3;

Db 273 QSKLEIIDAISELPFELGGSCCTADNCGCWRSDKGPWNPDPIMKRNVNGDHICSK---RS 329

Qy 248 QVKQQYEHSHVQISRGSSHOVE 268
| : ||::||
Db 330 QADNAGENI--ISQGNNSAVE 348

RESULT 11
F84539
hypothetical protein At2g16380 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84539
R.;Lin, X.; Kaul, S.; Rounsley, S.D.; Shen, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Lin, X.; Kaul, S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84539
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <STO>
A:Cross-references: UNIPROT:Q9SIW3; UNIPARC:UPI0000017A75C; GB:AE002093; NID:g4544391; PT:
C:Genetics:
A:Gene: At2g16380
A:Map position: 2

Query Match 10.8%; Score 218; DB 2; Length 582;
Best Local Similarity 29.2%; Pred. No. 2.2e-10;
Matches 59; Conservative 38; Mismatches 91; Indels 14; Gaps 3;

Qy 34 MWSESK-----RTLTTSLAWQPPEVIOQYLSGGCMGYDLDCGPVWDIIIGPKDAKGL 86
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 104 MWSMLQWRMDFGVDTIIEDFEFEIDQVLKHYPQGVHGVDKSRPVYIERLGQIDANKL 163

Qy 87 LFSASKQDRLTKMRCELLLQ----ECAHQTTKLGRKVETITIIYDCBGLGLKHLWKPA 142
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 164 LQATTMDRYEKYHVKEFKMKIFPPSCSAAKK--HIDQSTTFIDVGQVGLKNFNKSA 220

Qy 143 VEAYGEFLCPMEENYPETIKELFVVVAPKPPVAVYNLIKPLSDETRKKIMVLGANYPEV 202
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 221 RELQLRLKIDNDNYPETLNRMFIINAGPGFLLWAPIKFPDPTKTSKHVLGNKYQPK 280

Qy 203 LLKHISPDOVPVEYGGTMTDPD 224
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 281 LLEAIDASELPYFFGLCTCAD 302

RESULT 12
E84602
hypothetical protein At2g21540 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84602
R.;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84602
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <STO>
A:Cross-references: UNIPROT:Q9SIK5; UNIPARC:UPI00000A24B9; GB:AE002093; NID:g4567235; PT:
C:Genetics:
A:Gene: At2g21540
A:Map position: 2

Query Match 10.8%; Score 217.5; DB 2; Length 371;
Best Local Similarity 28.3%; Pred. No. 1.3e-10;
Matches 64; Conservative 35; Mismatches 96; Indels 31; Gaps 5;

